AD_____

Award Number: DAMD17-00-1-0227

TITLE: Mechanisms of Intraductal Tumor Spread

PRINCIPAL INVESTIGATOR: Carlos Ortiz de Solorzano, Ph.D.

CONTRACTING ORGANIZATION: University of California at Berkeley
Berkeley, California 94720

REPORT DATE: August 2003

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

# REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188



| 1. AGENCY USE ONLY (Leave blank) | 2. REPORT DATE | 3. REPORT TYPE AND DATES COVERED |
|---|---|---|
| | August 2003 | Annual (1 Aug 2002 - 31 Jul 2003) |

**4. TITLE AND SUBTITLE**
Mechanisms of Intraductal Tumor Spread

**5. FUNDING NUMBERS**
DAMD17-00-1-0227

**6. AUTHOR(S)**
Carlos Ortiz de Solorzano, Ph.D.

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
University of California at Berkeley
Berkeley, California 94720

E-Mail: codesolorzano@lbl.gov

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**

**11. SUPPLEMENTARY NOTES**

**12a. DISTRIBUTION / AVAILABILITY STATEMENT**
Approved for Public Release; Distribution Unlimited

**12b. DISTRIBUTION CODE**

**13. ABSTRACT (Maximum 200 Words)**
During the administrative funding period of this grant, we have developed a system that permits three-dimensional reconstruction of intraductal tumors (DCIS) from physical tissue sections. The system reduces the interaction required for low-resolution imaging of H&E stained sections, registration of images of consecutive sections and annotation of tissue structures -i.e. morphologically normal ducts and intraductal tumors-in the images. In addition, we have developed fully automatic tools for image registration and annotation that are now being integrated in our system and used in the reconstruction of the latest tissue specimens. Complementing morphological H&E based reconstructon, our system can be used for morphologically driven acquisition of high-resolution images from immunostained intermediate sections, both using fluorescence and brightfield microscopy. We are using our system to characterize the cellular differences between malignant transformed cells and morphologically normal cells of ducts either in continuum or in the proximity of DCIS tumors. Under the no-cost extension requested, we expect to complete the characterization, which includes the following markers: Her2, ER and PR status Ki67 (proliferation), Caspase 3 (apoptosis) and a marker of dedifferentiation of stemness, CD49f.

**14. SUBJECT TERMS**
Three-dimensional microscopy, computer aided tissue reconstruction, DCIS, FISH

**15. NUMBER OF PAGES**
64

**16. PRICE CODE**

| 17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
|---|---|---|---|
| Unclassified | Unclassified | Unclassified | Unlimited |

**Table of Contents**

## INTRODUCTION

When a cancerous lump is detected early in the breast, the patient may elect to undergo the less traumatic treatment of lumpectomy. The treatment involves removing the cancerous lesion while leaving most of the breast intact. Although the surfaces of the excised tissue are normally checked for signs of cancer to ensure that the lesion was completely removed, approximately 1 in 5 of these patients suffer from recurrence of the disease. Therefore, lumpectomy alone frequently does not render the patient disease free. We hypothesize that diseased cells exist outside the histologically identifiable border of the biopsied lesion, most likely in the form of individual or small groups of cells that have extended from the primary lesion. These cells could eventually create new secondary cancer foci. The alternative hypothesis is that the disease is really multicentric and exists as a system of independent non-connected (neither physically nor genetically) foci. This project uses 3D digital microscopy for analyzing tissue structure at multiple scales and in situ genetic analysis to recognize normal from diseased cells on an individual basis. By combining these two techniques, we will be able to measure the spatial distribution of genetically aberrant cells versus normal appearing cells beyond the leading edge of the intraductal lesions. These cancer cells lying in the lumen of morphologically normal ducts could be easily overlooked using traditional histology staining. The distribution of cancer cells will help us understand the spreading mechanism. Answering this question may help us to predict before surgery which patients will suffer recurrence, which patients need additional treatment following lumpectomy to avoid recurrence, and provide valuable information in the search for new treatments.

In this project we will use computerized microscopy to locate a lesion inside a duct and to trace in 3 dimensions the ducts branching from it. Then, using the technique described above we will detect the abnormal cells in the normal looking ducts extending from the lesion. Analysis of the spatial pattern of abnormal cells will tell us if isolated or small groups of cancer cells are present or if some other spreading mechanism is taking place, and how far the spreading is from the lesion.

**BODY**

Our accomplishments during this year (08/01/01-07/31/02) will be described following the Tasks enumerated in the approved proposal. Those tasks are listed below, and the sub-tasks corresponding to the final funding year have been underlined. Tasks completed during the first two years, including those originally scheduled for the second or third year, are highlighted in italic.

As explained in the previous two reports, there is an accumulated delay in the accomplishment of the proposed goals. Reasons for this delay are: 1) a three month delay in the release of funds for the first year of work, that in fact began in October 2000 rather than July 2000; 2) the change in tissue source described in the first year report; 2) the difficulty to find the desired transition from normal to intraductal carcinoma in the new tissue source, as described in the second year report; 3) the unexpected extra work and technical developments necessary for speeding up the acquisition, annotation and reconstruction of the tissue samples, as well for dealing with extremely large images of entire tissue sections, as described in both first and second year reports; 4) an approved six month leave of absence of the Pi during the second year; 5) the delay in the development and integration of cell-level analysis tools, due in part to the poor performance of the postdoctoral fellow hired for that task, who left the lab after one year without having provided the desired software.

In consequence, in spite of to the work done to speed-up the tissue processing and analysis, we have not completed all the aims of the original proposal. Accordingly, we have requested a no-cost one-year extension of the grant that will provide enough time to complete the goals. The funding for this year is secured by carried-forward money that has not been spent during the administrative length of the grant.

As mentioned in last year's report, the technology developments required for this grant are being done under the join budget of this and our other grant, "*Three-dimensional computer-based mammary gland reconstruction for measurement of the patterns of hormone receptor expression during mammary development*" (DAMD17-00-1-0306). That is why the technology accomplishments reported here (mainly Task1) are similar to those described in the other grant's report.

**Task 1**. (Months 1-12) Modify an existing microscopic imaging system for acquiring low magnification (1 pixel= 5 μm) images of entire tissue sections and for tracing in 3D the ducts in the tissue specimen from a series of images of adjacent sections.

1. Complete the existing JAVA based software for interactive marking and 3D virtual rendering of ducts so that it allows any branching pattern. (Months 1-6)
2. Interface the existing acquisition and registration software with the JAVA application to allow revisiting of acquired slides for inspection and high-resolution acquisition of areas of interest. (Months 6-12)

At the end of the entire funding period, we can present R3D2, a robust JAVA based software that can be used to semi-automatically image and reconstruct tissue structures from serially sectioned thin (5 μm) tissue sections. This system controls a fully automated scanning microscope and can automatically acquire entire sections stained for either fluorescence (e.g. DAPI) or bright field (e.g. H&E) microscopy. The system acquires and tiles together the multiple single field-of-view images that cover the entire extent of the section, adjusting the focus plane of the microscope, when required to maintain optimum contrast of the acquired images. All the related sections that make up a tissue block are stored following a predefined directory structure and can be loaded and browsed easily. To allow real-time loading and browsing of these extremely large images (some of them can reach 70Mb) we have used memory mapping techniques that permit accessing the large image files without loading the entire image pointer in the computer memory. Visualization of the entire section is done by displaying reduced version of the original imaged, subsampled to fit the size of the visualization window. Then, zooming in areas of the section is done by retrieving the image data directly from the image file, based on image coordinates selected in the subsampled version of the image of the entire section.

Image acquisition can be done in grayscale using the 12 bit CCD camera attached to the microscope, or in color by sequentially acquiring grayscale images using the proper excitation/emission filters (fluorescence) or an RGB tunable liquid crystal filter (brightfield). Visualization of 12 bit images in a 8 bit per color depth display can be done by linear compression or truncation of the image data. In addition, images can be linearly stretched to enhance low contrast images.

To complement its tissue imaging capabilities, R3D2 provides interactive tools for registering images of consecutive sections, which is necessary to faithfully render in

three dimensions tissue structures traversing multiple sections. This is critical, due to the significant missregistration between sections, consequence of the entire manual sectioning process. In fact, besides linear effects (shifting and rotation), there are frequent non-linear effects such as tissue folding, shrinking, stretching or tearing which make registration extremely difficult. Within the scope of this grant, we have solved and automated (see bellow) the linear registration problem, but correcting non-linear effects is one of the future tasks that we would like to address in the near future.

R3D2 is equipped with annotation tools (active pencil, selection, grouping, splitting, etc) that can be used to manually delineate and connect (within and between sections) tissue structures (in our application normal ducts and lobulo-alveolar structures, DCIS tumors, areas of invasive carcinoma, etc), that can be rendered in 3D using a modified Delaunay triangulation and the surface rendering OpenGL-based toolkit embedded in Java3D. As shown, Java ensures seamless continuity between image acquisition, annotation and reconstruction (rendering). The 3D rendering of the tissue is interactive, in that the entire tissue "scene" can be seen from different view points, at the distance and angle of the user's choice. Also, specific tissue "volumes" can be selected to retrieve information.

Finally, R3D2 allows revisiting the original tissue slides from both the images of the entire sections or from the 3D reconstruction of the tissue. Revisiting can be uses to visually inspect the slide under the same of different optical conditions (i.e. lense magnification, light-filtering, etc) although is normally used in to acquire multiple-color areas of interest at high magnification.

A detailed description of the system, extending what has been summarized above, has been published in the journal "Microscopy Research and Technique" (Bibliography J1 –Appendix 1-), and before publishing, presented in International conferences (Bibliography P1, A1, A2, A3). Some news agencies covered and distributed new releases that were published in several newspapers and online news services (Bibliography O1, O2, O3, O4) .

Following the described developments, and in the process of applying R3D2 to the reconstruction of DCIS specimens, it became clear that new technical developments were necessary to increase the throughput of the system, since the time required to reconstruct one case using the existing mostly manual tools was in the order of two months. The two main bottlenecks in the process described above are the manual registration and annotation of the sections. Therefore we developed automated tools for both tasks, which have been integrated in the system and are currently being used, providing substantial time savings.

We developed a multiscale, multiresolution registration algorithm based on gradient correlation between consecutive image sections. The algorithm calculates the optimum *rigid body* transformation (rotation plus translation) between each two consecutive images. To reduce the computational cost, we start using heavily subsampled images, and refining the registration using decreasingly subsampled images. At each iteration level, the registration is calculated as follows: first, the gradient of both images (reference and registering) is calculated and thresholded using an adaptive threshold. Then the distance transform of the images is obtained, that contains the distance of each point to the nearest gradient area (i.e. boundary). Finally the distance transform corresponding to the image being registered is scanned over that of the reference image, for different rotations and translations, and the optimum registration is defined as the absolute minimum of the product of both images, which ideally corresponds to 0, i.e., to perfect overlap between both sections. A more detailed description of the algorithm, including examples has been recently submitted and accepted for presentation at an international conference (Bibliography A4)

To address the second bottleneck, which is the annotation of histological structures, we paid attention to the increasing interest in the application of partial differential equation (PDE) morphologically driven flows (i.e. Level Set methods) in image processing and analysis. A description of the Level Set (LS) methodology is out of the scope of this report. Therefore, a very succinct user-focused is provided next. In a nutshell, the LS approach considers the image as a force or energy field determined by one or a combination of selected image features (e.g. intensity, gradient, object curvature, distance...). Then the segmentation of objects is done by letting some initial seeds manually placed on the original image evolve under the driving force of a velocity function that depends on the energy field. This way, assumed that the right energy field

is selected, the curves (surfaces in 3D) that define boundaries of the seeds will converge in or near the boundaries of the objects that one wants to extract.

The initial curve is represented here as the zero level set of a higher dimensional function, and the motion of the curve is embedded within the motion of that higher dimensional function. The speed of that motion is defined based on the characteristics of the image to be segmented. In our case the speed is adjusted so that when the interface is on top of areas with a low gradient it expands quickly, whereas when the gradient is large (indicating the location of an edge in the image) the curve is slowed down. In addition, a surface tension term is included in the speed function to slightly retard or accelerate the contour depending on its curvature, thereby preserving the smoothness of the advancing front. This approach offers several advantages. First, the zero level set of the higher dimensional function is allowed to change topology and form sharp corners. Second, geometric quantities such as normal and curvature are easy to extract from the hypersurface. Finally, everything expands directly to three dimensions if we embed the advancing three-dimensional surface as the zero level set of a four-dimensional function.

However, changing an n-dimensional problem into one in n+1 dimensions increases the computational cost associated with the method. The narrow-band approach accelerates the level set flow by updating the position of the curve only in a narrow vicinity of its current location. But in our experience, the narrow band technique does not reduce the computational cost to a reasonable limit, due to the large size of the images of the sections. Thus, we propose to use the fast marching method, a numerical technique to solve the equation that drives the movement of the curve by combining an efficient –constrained- solution to the equation of the movement of the front, narrow-band level set methods and a min-heap data structure. This method is only used for monotonically advancing fronts (speed always positive or negative), providing a result very fast, albeit not as accurate as the one obtained by using the level set algorithm. This result is then used as the initial condition for the slower but more accurate level set segmentation.

We have used this approach to segment histological structures on H&E and fluorescent (counterstained) sections, by placing seed points either in the background or in the lumen. The results have been published this year in the peer reviewed conference proceedings of the SPIE Biomedical Optics Conference in San Jose, CA, (Bibliography P2 –Appendix 2-) and a full paper describing the method and results has been submitted to the Journal of Biomedical Optics. (Bibliography J3 –Appendix 3-)

**Task 2.** (Months 1-30) Using invasive cancer specimens with intraductal extension of DCIS, identify DNA loci that are amplified in a high proportion of the cells of the invasive lesion using CGH.

1. Select 9 mastectomy specimens following the criteria described in the Methods section of the Proposal body (Months 1-6)
2. Section and H&E stain mastectomy specimens (Months 6-12, 2 specimens; Months 12-24, 5 specimens; Months 24-30, 2 specimens)
3. Acquire sections using our registration software (Months 12-24 6 specimens; Months 24-30, 3 specimens)
4. Reconstruct the mammary ducts and identify the leading edge of the intraductal component (Months 12-24, 5 specimens; Months 24-30, 4 specimens)

As it was mentioned in the report written at the end of the first year, the mastectomy tissue originally reserved for this study was not preserved properly. Due to this fact, we looked for a new source of tissue that was found at the UCSF Comprehensive Cancer Center's tissue repository. However, instead of full mastectomies where the first task would have been to look macroscopically for areas likely to show intraductal extension of DCIS, we were given access to paraffin-embedded archived material. We received specimens four (UCSF1, UCSF2, UCSF3, UCSF4), consisting on one isolated tissue block per case, with the corresponding histological evaluation and Her2 status, done on the top slide of the block. Although the fact that none of the blocks showed a transition from DCIS to normal tissue on the top slide, they all had areas of normal tissue and DCIS. This led us to think that they could have had the desired transition. In fact, none of them had it. Fig 1 and 2 show examples of reconstructions of two of the cases (UCSF1 and UCSF2).

As described in the report corresponding to the second year, we continued looking for tissue having the desired normal epithelium to DCIS transition. We asked Dr. Alexander Borowsky from the UC Davis Pathology Department and Center for Comparative Medicine, who gladly identified several tissue specimens from his Department, and helped us identifying those areas. We are very greatful to him, since we finally have access to the tissue needed for this grant. We have received one case (UCD1) that is almost completely processed, as far as what this task is concerned, i.e, imaged at low resolution and reconstructed in 3D (at least the areas that have the

transition). During the one year no cost extension requested we would be able to study 4 other cases to reach the 9 mentioned in the original plan.

**Task 3.** (Months 1-30) Use FISH with probes to the loci identified by CGH

1. Do CGH on the selected mastectomy specimens (Months 1-6)
2. Do FISH to the two most amplified regions and a normal part of the genome (for control purposes) on intermediate sections to those used for the reconstruction of the ductal system (Months 6-12, 2 specimens; Months 12-24, 5 specimens; Months 24-30, 2 specimens)

The reason for using CGH was to be able to identify an amplified chromosomal locus that could be used to track pre-malignant, but otherwise morphologically normal cells that could move ahead of the leading edge of the intraductal invasion. That option, which would had been possible due to the abundance of tissue, had we used the original mastectomy specimens, was not feasible with the single paraffin-embedded tissue blocks provided by the UCSF Cancer Center Tissue Repository, therefore all the chosen cases were selected being Her2+, and we used the amplification of that oncogen as a possible marker to detect isolated, premalignant cells. FISH to the Her2 gene, located in chromosome 17, and a centromeric probe for ctr. 17 (control) was done in cases UCSF1, UCSF2 and UCSF3.

UCD1, due to reasons that will be explained below, was immunostained for the following proteins: Ki67 (proliferation), Caspase3 (apoptosis), Estrogen Receptor, Progesterone Receptor, β1 integrin (postulated as a marker for cells with a progenitor phenotype), CD49f (possible marker for mammary gland stem cells). UCSF4 will be immunostained with antibodies to the same set of proteins as the Case from UCD1, and so will be the four cases that we plan on doing during the one-year no-cost extension of the grant.

**Task 4.** (Months 18-36) Use high magnification (1 pixel= 0.5 μm) fluorescence microscopy, look for individual cells with the same amplified loci in the ducts emanating from the DCIS lesions. Assuming that we see genetically aberrant cells, measure the spatial relationship of these cells to the surrounding cells in order to characterize the pattern of aberrant cells and thus to provide information about the spreading mechanism of the disease.

1. Automatically enumerate FISH spots and measure the spatial distribution of aberrant cells (if found) in the histologically normal ducts starting at the very front of the intraductal tumor expansion. (Months 18-24, 4 specimens; Months 24-36, 5 specimens)

We proposed to use FISH to an amplified chromosomal locus on the tissue blokcs, and use the amplification as a marker of malignancy that could be tracked in normal ducts connected to the DCIS lesions. We hypothesized that some abnormal isolated cells could be the front of the advancing DCIS front. These isolated cells would be easily missed by routine pathological analysis of the tissue. An alternative hypothesis, the genetic transformation of the cell could be the consequence of a field-effect, and could precede any morphological transformation. As described in Task 2, the transition was not found in any of the first four cases received (UCSF1, UCSF2, UCSF3, UCSF4). However, we considered the quantification of the amplified gene worth doing, to compare the number of copies of the gene in DCIS areas with those of nearby normal ducts since, although unlikely those areas might be in fact connected in the rest of the tissue, outside the block.

Although the analysis is still underway, so far we have not found abnormal, amplified cells in normal ducts and alveoli located near DCIS tumors. This works against our hypothesis, saved the fact that we don't know for certain that both types of tissue are not connected outside the limits of the block. Another consideration to mention here is that the efficiency of the hybridization seems to depend on the type of tissue, and it is very difficult to optimize it to work well for both normal and tumor areas. The Cytogenetics Core of the UCSF Cancer Center, which performed the hybridization, provided staining that was preferentially optimized for tumor areas. In consequence some normal areas of all tissue blocks presented no signals.

Another argument against the use of FISH to a single amplified gene as a marker of transformation is that genetic markers found in a malignant cancer are not necessarily present in preneoplastic lesions, as a consequence of the clonal progression of the disease, which evolves by replacing complete populations of cells by new, mutated cells with a proliferative advantage.

These two problems, plus a methodological one having to do with the difficulty of fluorescence-hybridizing and imaging a high number of slides (up to 100) under similar conditions and without suffering the effects of fading, led us to look for an alternative approach, that we are using in UCSF4 and UCD1 and we will use in the four cases last cases. The alternative takes advantage of the number of consecutive sections to use several markers, instead of limiting ourselves to a single marker. The markers that we use are not genetic but epigenetic. Some of them are global or phenotypic, and therefore are not related to a particular genetic event. This way we can still look at preneoplastic changes in the normal areas connected of surrounding the DCIS tumors, by combining the expression of several markers. We are using markers of proliferation and apoptosis (ki67, caspase3); others are markers of lack of differentiation (β1, CD49f), which can reveal the presence and distribution of self-renewing cells with stem cell characteristics responsible for the rupture of the proliferative equilibrium of the gland; finally we are using markers of differentiation and hormonal responsiveness (PR, ER) that will be seen combined to the markers of 'stemness'. We are now in the process of imaging the immunostained sections corresponding to several transition areas found in UCD1. At the same time we are now immunostaining UCSF4 and we will do the same with the future four cases that we will analyze during the one year extension of the grant.

## KEY RESEARCH ACCOMPLISHMENTS

The main accomplishments achieved from this project are listed below. They will be completed during the one-year extension requested for this project, justified in page 2 of this report.

- We have developed a system that allows semi-automatic 3D reconstruction of tissue samples from fully sectioned tissue blocks. The system allows acquisition of entire tissue sections at low magnification, both in bright field and fluorescence microscope, registration of images of consecutive sections and annotation and 3D rendering of tissue structures (epithelial, endothelial, etc). The system allows revisting of areas of the tissue at higher magnification from both the 3D reconstruction of the tissue as wells as from the low resolution images of the sections. (Bibliography J1 –Appendix 1-), P1, A1, A2, A3, O1, O2, O3, O4)

- To solve the bottlenecks of tissue imaging and reconstruction, we have developed tools that automatically register consecutive sections (Bibliography A4) and for the unmanned segmentation of histological structures using geometrically driven image flows (Bibliography P2 –Appendix 2-, J3 –Appendix 3-)

- We have developed software that can segment counterstained nuclei and quantify FISH signals or gene expression from the fluorescent high-magnification areas of interest.

- So far we have successfully imaged, reconstructed and revisited at high resolution five biopsy of tissue from a patient with Ductal Carcinoma Is Situ (DCIS) of the breast. (UCSF1, UCSF2, UCSF3, UCSF4, UCD1). The tissue was fully sectioned and alternatively stained with H&E and a nuclear fluoresecent counterstain (DAPI) plus FISH with a probe against the DNA locus of the *erb-b2* producing gene (UCSF1, UCSF2, UCSF3). UCD1, which presents several areas of transition normal-DCIS was also alternatively stained with H&E and immunostained for the following markers (ki67, caspase3, β1 integrin, CD49f, ER, PR). UCSF4 will be stained as UCD1.

- The analysis of UCSF1, UCSF2, UCSF3 performed so far does not reveal the existence of transformed cells within the normal epithelium neighboring DCIS areas. However, as explained above, it is likely that there is no physical connection between the normal ducts and DCIS tumors, in which case we would not expect to see a transformation. We continue the analysis, as described, using UCSF, UCD1 and the future 4 cases that we will received during next year.

## REPORTABLE OUTCOMES

Manuscripts (published or accepted):

- "*A system for combined three-dimensional morphological and molecular analysis of thick tissue samples*" Fernandez-Gonzalez R., Jones A., Garcia-Rodriguez E., Chen P.Y., Idica A., Barcellos-Hoff M.H., Ortiz de Solorzano C. *Microscopy Research and Technique* 59(6):522-530, 2002..
- "*Recent advances in quantitative digital image analysis and applications in Breast Cancer*". Ortiz de Solorzano C., Callahan D.E., Parvin B., Costes S., Barcellos-Hoff, M.H. *Microscopy Research and Technique* 59(2):119-127, 2002.
- "*A geometric model for image analysis in cytology*" Ortiz de Solorzano C., R. Malladi, Lockett S. In: Geometric methods in bio-medical image processing. Ravikanth Malladi (Ed.). Springer Verlag 2002, pp. 19-42.
- *"Automatic segmentation of histological structures in mammary gland tissue sections". Fernandez-Gonzalez R., Deschamps T., Idica A.K., Malladi R., Ortiz de Solorzano C. Accepted for the special issue on Women's Health of the Journal of Biomedical Optics.*

Manuscripts (in preparation):

- "*Three-dimensional Histo-pathology of the mammary gland*" To be submitted to *The Journal of Mammary Gland Biology and Neoplasia*

Conference Proceedings:

- "*A system for computer-based reconstruction of 3-dimensional structures from serial tissue sections: an application to the study of normal and neoplastic mammary gland biology*". Fernandez-Gonzalez R., Jones A., Garcia-Rodriguez E., Knowles D., Sudar D., Ortiz de Solorzano C. Proceedings Microscopy and Microanalysis'01. Microscopy and Microanalysis 7, Supplement 2, pp.964-965, 2001
- "*Automatic segmentation of structures in normal and neoplastic mammary gland tissue sections*". Fernandez-Gonzalez R., Deschamps T., Idica A.K., Malladi R., Ortiz de Solorzano C., Proceedings of Photonics West 2003, Vol. 4964, 2003

Presentations:

- *A system for computer-based reconstruction of 3-dimensional structures from serial tissue sections: an application to the study of normal and neoplastic mammary gland biology.* Microscopy and Microanalysis'01, Long Beach, CA August $5^{th}$-$9^{th}$, 2001. Platform presentation.
- "*3D Histo-Pathology: towards a morphological characterization of ductal carcinoma in situ of the breast*" Annual Meeting of the American Association for Cancer Research (AACR). San Francisco, CA, April 4-9, 2002.
- *"3D Mammary Histopathology"* Fernandez-Gonzalez R., Idica A. K., Ortiz de Solorzano C. 2003 Mammary Gland Biology Gordon Research Conference . Roger Williams University, Bristol, Rhode Island, June 1-6, 2003.
- *"Automatic Registration of Mammary Gland Section Images"* Arganda-Carreras I., Fernández-González R., Ortiz de Solórzano C. First International Meeting on Applied Physics (APHYS 2003), Badajoz, Spain, October 13th-18th, 2003

Informatics:

- As described in the Body of the report and in the Reportable Outcomes sections, we have developed and integrated new methods to automatically extract histological information from tissue sections, as well as morphological and molecular information at the cellular level.

Funding obtained:

- *Segmentation of Mammary Gland Ductal Structure Using Geometric Methods.* P.I.'s Malladi R. and Ortiz de Solorzano C. Granted by the LBNL Laboratory Directed Research and Development Program (LDRD), in the Strategic-Computational Sub-Program. Period Oct 2001- Sept 2004
- *Characterization of Adult Stem Cell Involvement in Mammary Gland Development.* PI: Dr. Carlos Ortiz de Solorzano Funded by: LBNL Laboratory Directed Research and Development Program (LDRD). Period Oct 2002-Sept 2004
- *Three-dimensional Modeling of breast cancer progression.* PI: Dr. Carlos Ortiz de Solorzano. Funded by: University of California, Breast Cancer Research Program Grant Number – 8WB-0150

- *"Characterization of label-retaining cells and their niche in the mouse mammary gland"* DOD Breast Cancer Research Program Predoctoral Fellowship. PI. Rodrigo Fernández-González.
- NASA NSCOR Program Project. PI. Mary Helen Barcellos-Hoff, Ph.D.

Funding applied:

- "*An automated system for three-dimensional Histopathology*". NIH R21/R33 Phase Innovator Award, for the PAR: "Technology developments for Biomedical Applications"/ Submitted on June 1st, 2003. PI. Carlos Ortiz de Solórzano, Ph.D.
- "*Biological Basis and Functional Phenotypes of Breast Density*". NIH Program Project. PI: Thea Tlsty, Ph.D.

Employment or Research:

- Due to the successful performance of the PI as a Scientist during the first two years of the project, he has been promoted to a Staff Scientist Position at the Life Sciences Division, Lawrence Berkeley National Laboratory of the University of California.
- This grant has partially supported Mr. Rodrigo Fernández-González, a Ph.D. candidate in the joint UC Berkeley-UC San Francisco Program in Bioengineering. Rodrigo continues working with me part time as a Graduate Student Research Assistant. On January 2003 he was granted a DOD-BCRP predoctoral fellowship *"Characterization of label-retaining cells and their niche in the mouse mammary gland"* that will support him until the end of his graduate work.
- Half way through the reporting period (in January 2002), Dr. Umesh Adiga, a Ph.D. in Computer Sciences, joined my lab as a postdoctoral fellow to work on the image analysis required for the automatic segmentation of nuclei and FISH signals, as well to other image analysis and processing tools required for this project. He left the laboratory in January 2003 due to his poor performance.
- In November 2003, Dr. Ouahiba Laribi joined our lab. She holds a recent Ph.D. in Molecular and Cell Biology. Her part in this grant is relatively small, and mainly has to do with the immunostaining and sample preparation.
- Mr. Adam Idica, an Integrated Biology undergraduate student at UC Berkeley has been working in this project for two years. He recently graduated from Berkeley and has decided to continue working in our lab as Technical Assistant.

- **CONCLUSIONS**

In summary, during the administrative length of the project, we have developed the computer and microscopy platform that we proposed to develop. The developments have being slower than expected due to the realization of the need for automating some of the time consuming tasks, namely the registration and annotation of the images of the sections. The developments have been or are in the process of being published in peer-reviewed journals and have been successfully presented in international conferences. We believe that this advanced computerized microscopy platform, developed thanks to the funding support of this grant, will be of use in many future studies that required looking at molecular events at cellular level within the tissue context where they occur.

Due to the above mentioned delay in the technical developments, which was accompanied by other factors listed in pg.2 of this report, the analysis, which aims at locating signs of preneoplastic transformation in normal epithelium connected to DCIS lesions, has suffered a significant delay. Out of the 9 proposed samples, we have processed 5 tissue samples. Although the analysis has not been completed yet, we have not found any signs of preneoplastic transformations in the normal epithelium of three of the five cases processed. Due to the reasons described in the description of Task4, we have modified the original plan, and in the two cases that we have not analyzed yet we will look at multiple phenotypic markers instead of using a single genetic marker (amplification of Her2). We expect that this approach is more appropriate to detect symptomatic changes, which don't have to be necessarily associated to a particular marker. We will use the same approach for the four cases that we will process during the requested one year no-cost extension of the project.

## BIBLIOGRAPHY


Papers in Peer Reviewed Journals:

**J1. A system for combined three-dimensional morphological and molecular analysis of thick tissue samples**. Fernandez-Gonzalez R., Jones A., Garcia-Rodriguez E., Chen P.Y., Idica A., Barcellos-Hoff M.H., Ortiz de Solorzano C. *Microscopy Research and Technique* 59(6):522-530, 2002.

**J2. Applications of quantitative digital image analysis to breast cancer research**. Ortiz de Solorzano C., Costes S., Callahan D.E., Parvin B., Barcellos-Hoff M.H. *Microscopy Research and Technique* 59(2):119-127, 2002

**J3. Automatic segmentation of histological structures in mammary gland tissue sections.** Fernandez-Gonzalez R., Deschamps T., Idica A.K., Malladi R., Ortiz de Solorzano C. *Accepted for the special issue on Women's Health of the Journal of Biomedical Optics*

Book chapters:

**B1. A geometric model for image analysis in cytology**. Ortiz de Solorzano C., R. Malladi, Lockett S. In: Geometric methods in bio-medical image processing. Ravikanth Malladi (Ed.). Springer Verlag 2002, pp. 19-42.

Conference Proceedings:

**P1. A system for computer-based reconstruction of 3-dimensional structures from serial tissue sections: an application to the study of normal and neoplastic mammary gland biology**. Fernandez-Gonzalez R., Jones A., Garcia-Rodriguez E., Knowles D., Sudar D., Ortiz de Solorzano C. Proceedings Microscopy and Microanalysis'01. Microscopy and Microanalysis 7, Supplement 2, pp.964-965, 2001

**P2. Automatic segmentation of structures in normal and neoplastic mammary gland tissue sections.** Fernandez-Gonzalez R., Deschamps T., Idica A.K., Malladi R., Ortiz de Solorzano C., Proceedings of Photonics West 2003, Vol. 4964, 2003

Abstracts:

**A1. Three-dimensional Characterization of Ductal Carcinoma In Situ of the Breast.** Carlos Ortiz de Solórzano, Rodrigo Fernández-González, Adam Idica, Fred Waldman, Joe W. Gray. Era of Hope Meeting. Orlando, FL. September 25-28, 2000.

**A2. 3D Histo-Pathology: towards a morphological characterization of ductal carcinoma in situ of the breast.** Carlos Ortiz de Solorzano, Rodrigo Fernandez-Gonzalez, Koei Chin, Karen L Chew, Fred M Waldman, Joe W Gray. Annual Meeting of the American Association for Cancer Research. San Francisco, CA April 2002.

**A3.** 3D Mammary Histopathology. Fernandez-Gonzalez R., Idica A. K., Ortiz de Solorzano C. 2003 Mammary Gland Biology Gordon Research Conference . Roger Williams University, Bristol, Rhode Island, June 1-6, 2003.

**A4.** Automatic Registration of Mammary Gland Section Images. Arganda-Carreras I., Fernández-González R., Ortiz de Solórzano C. First International Meeting on Applied Physics (APHYS 2003), Badajoz, Spain, October 13th-18th, 2003

Online references:

**O1.** Laser Focus World:
http://lfw.pennnet.com/Articles/Article_Display.cfm?Section=ARCHI&ARTICLE_ID=177510&VERSION_NUM=1&p=12

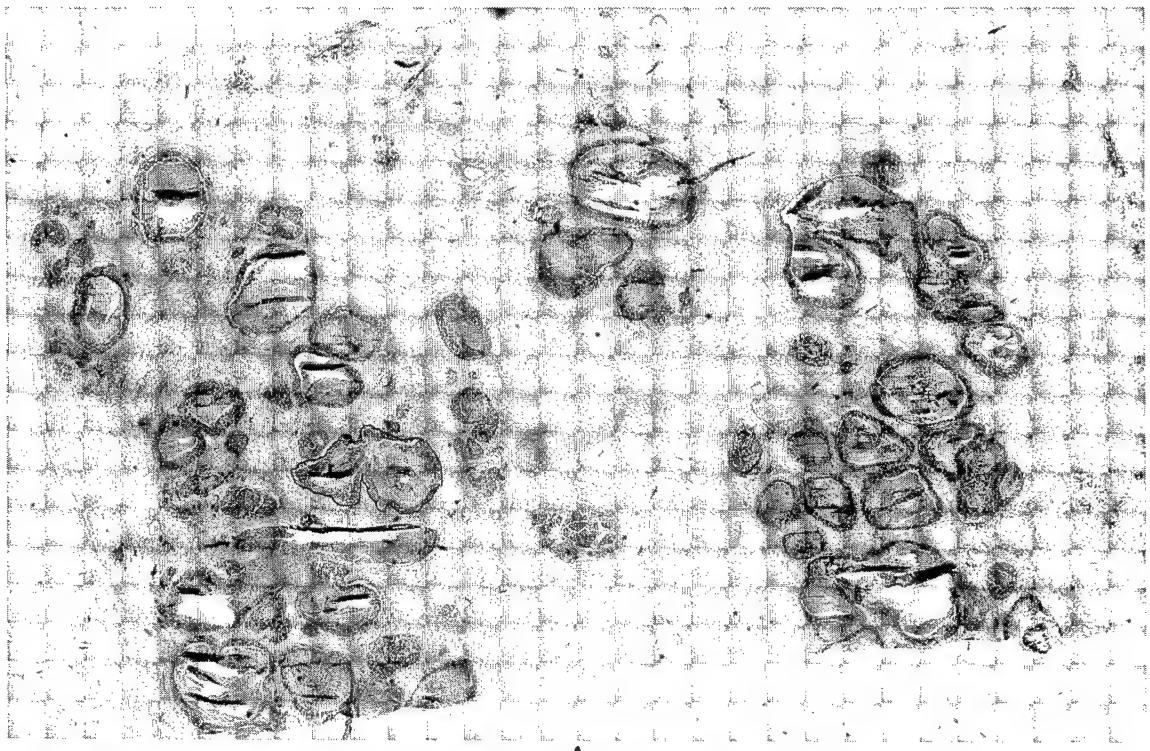
**O2.** Cipherwar Science new service: http://www.cipherwar.com/lists/lbnl/msg00025.html

**O3.** DailyCalifornian Science and Technology:
http://www.dailycal.org/article.asp?id=11347&ref=search

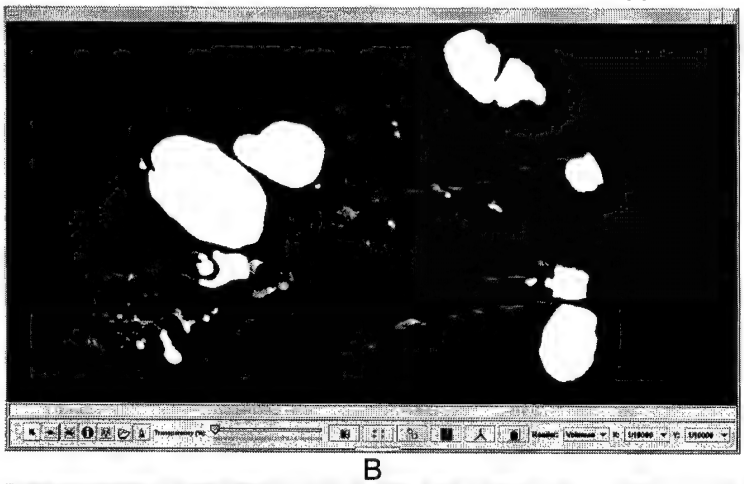
**O4**. Lawrence Berkeley National Laboratory New Release:
http://www.lbl.gov/Science-Articles/Archive/LSD-tracking-cancer-3D.html

**PERSONNEL**

- Carlos Ortiz de Solórzano, Ph.D. Principal Investigator
- Ouahiba Laribi, Ph.D. Postdoctoral Fellow (from 11/02)
- Umesh Adiga, Ph.D. Postdoctoral Fellow (from 01/02 to 01/03)
- Rodrigo Fernández-González, Graduate Student Research Assistant (from 10/00)
- Adam Idica, Technical Assistant (from 10/01)
- Shamroze Khan, Undergraduate Technical Assistant (from 05/03)



A

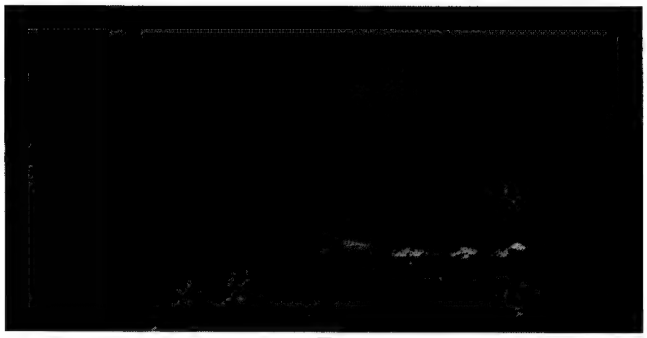


B

C

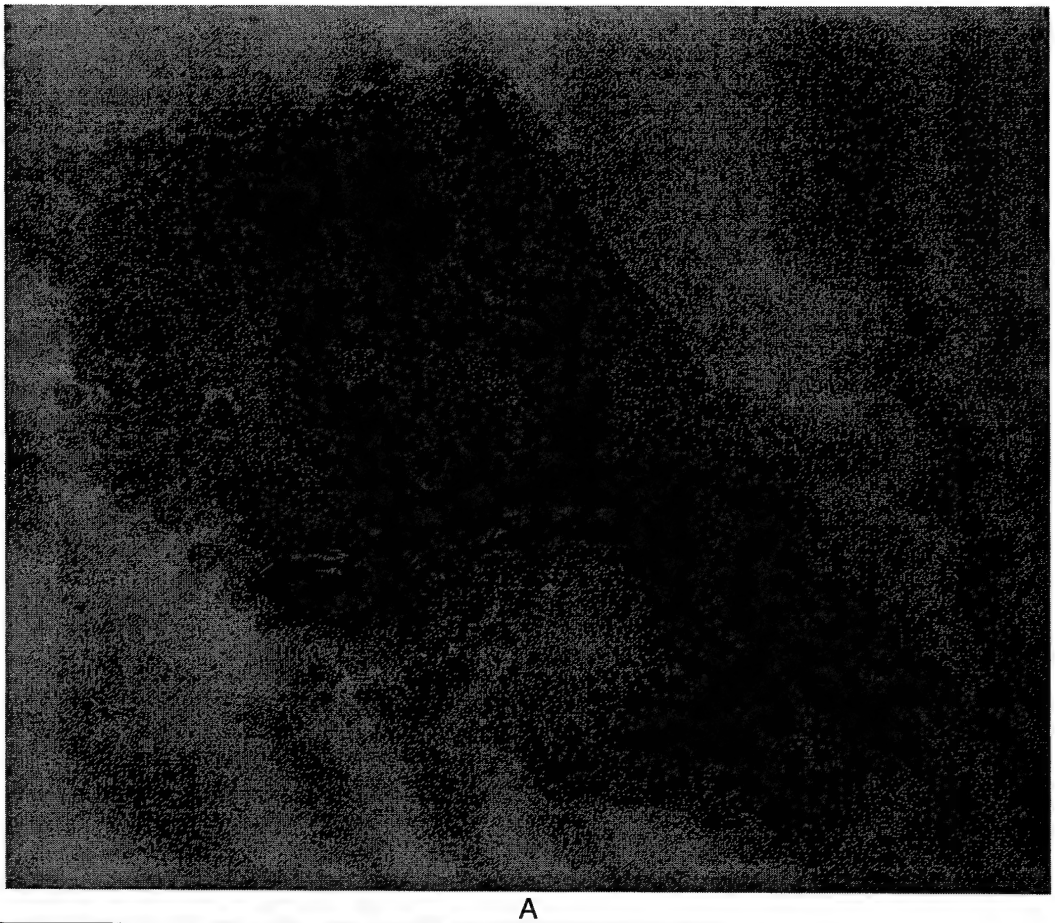
D

E

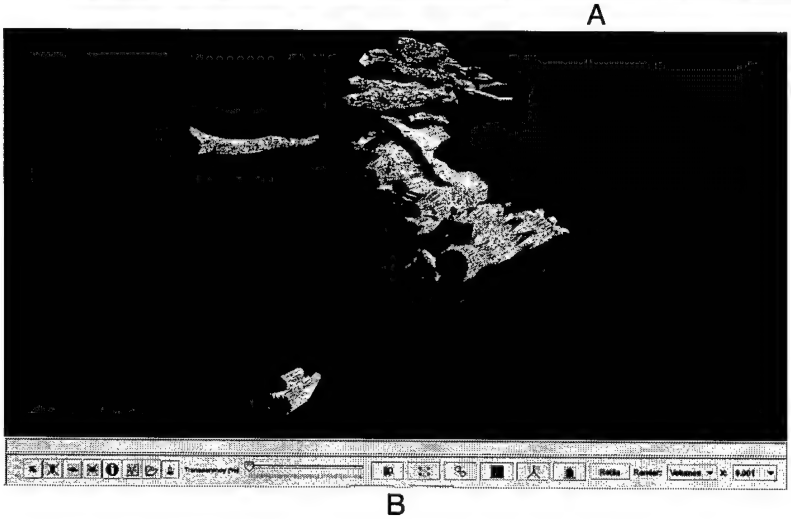


F

Figure 1. Reconstruction of case UCSF1. A) Low resolution (2.5X) image of an H&E stained section. B) 3D reconstruction of selected tumor bodies and normal areas. C) Reduced version of A, where red lines indicate the structures rendered in B. D) High resolution (40X) Area of a consecutive section, counterstained with DAPI. E) Same areas in D, hybridized with a FITC labeled centromeric probe of chr. 17. F) Same area as in D,E, hybridized with a CY3 labeled probe to the Her2 (erbb2) gene in chromosome 17.



A



B

C

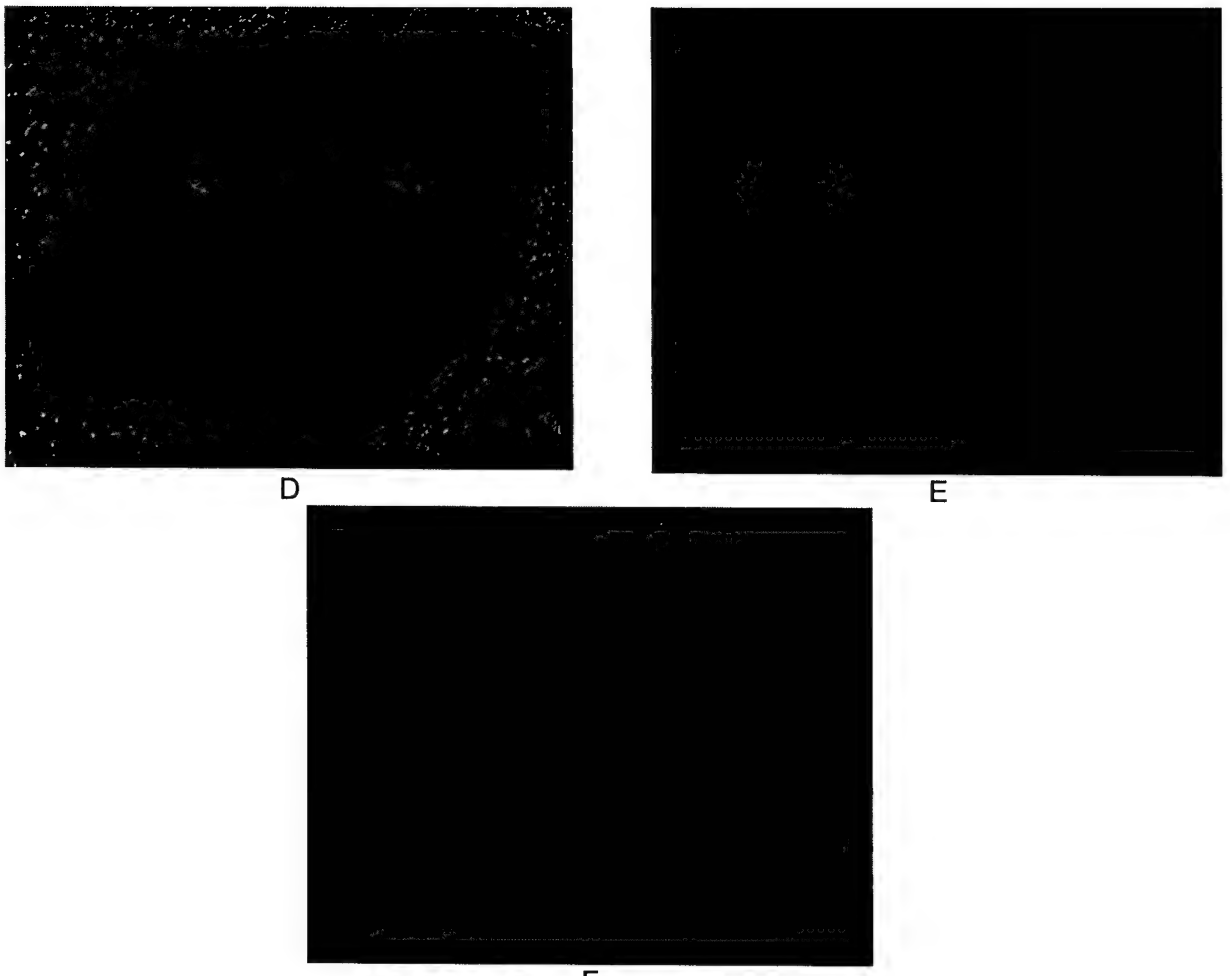



Figure 2. Reconstruction of case UCSF2. A) Low resolution (2.5X) image of an H&E stained section. B) 3D reconstruction of selected tumor bodies and normal areas. C) Reduced version of A, where red lines indicate the structures rendered in B. D) High resolution (40X) Area of a consecutive section, counterstained with DAPI. E) Same areas in D, hybridized with a FITC labeled centromeric probe of chr. 17. F) Same area as in D,E, hybridized with a CY3 labeled probe to the Her2 (erbb2) gene in chromosome 17.

# System for Combined Three-Dimensional Morphological and Molecular Analysis of Thick Tissue Specimens

RODRIGO FERNANDEZ-GONZALEZ, ARTHUR JONES, ENRIQUE GARCIA-RODRIGUEZ, PING YUAN CHEN, ADAM IDICA, STEPHEN J. LOCKETT, MARY HELEN BARCELLOS-HOFF, AND CARLOS ORTIZ-DE-SOLORZANO*

*Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720*

*KEY WORDS* computer assisted microscopy; 3D reconstruction; assisted histopathology; JAVA

*ABSTRACT* We present a new system for simultaneous morphological and molecular analysis of thick tissue samples. The system is composed of a computer-assisted microscope and a JAVA-based image display, analysis, and visualization program that allows acquisition, annotation, meaningful storage, three-dimensional reconstruction, and analysis of structures of interest in thick sectioned tissue specimens. We describe the system in detail and illustrate its use by imaging, reconstructing, and analyzing two complete tissue blocks that were differently processed and stained. One block was obtained from a ductal carcinoma in situ (DCIS) lumpectomy specimen and stained alternatively with Hematoxilyn and Eosin (H&E), and with a counterstain and fluorescence in situ hybridization (FISH) to the ERB-B2 gene. The second block contained a fully sectioned mammary gland of a mouse, stained for histology with H&E. We show how the system greatly reduces the amount of interaction required for the acquisition and analysis and is, therefore, suitable for studies that require morphologically driven, wide-scale (e.g., whole gland) analysis of complex tissue samples or cultures. *Microsc. Res. Tech. 59:522–530, 2002.* Published 2002 Wiley-Liss, Inc.†

## INTRODUCTION

Understanding complex biological systems requires tissue-level integration of information from multiple sources such as molecular, physiological, anatomical, and so on. However, none of the existing analytical methods in biology provides the required level of integration, in that they either do not account for intercellular variation (RFLP, Southern blots, microarray technologies, etc.), or they do it at the expense of tissue integrity (e.g., flow cytometry).

Image-based cytometry (IC) can provide molecular or genetic information (e.g., by using fluorescence in situ hybridization [FISH] or immunohistochemistry [IHQ]), in fixed cells within their native morphological tissue context. Volumetric, 3D morphological information can be obtained using confocal laser scanning microscopy (CLSM). However, only relatively thin tissue sections (<100 μm) can be studied, due to light scattering and refractive index mismatch problems that occur when imaging deeper in the tissue, and because of practical limitations of effectively staining thicker sections by IHQ or FISH.

When analysis and integration of molecular and morphological information at high resolution is aimed on a bigger scale (e.g., tissue or a small gland), the only existing approach requires sectioning the tissue, followed by both histological and molecular staining of consecutive tissue sections. The analysis is normally performed by visual inspection of the sections under the microscope. This approach greatly limits the extent and accuracy of the analysis, due to the difficulty that our visual system finds when composing (extrapolating) meaningful 3D information from a series of 2D sections. Since only a few colors (2–4) can be discriminated, both in bright field and fluorescence microscopy, along with other practical problems related to multicolor IHQ or (F)ISH, we have to do the histological and the molecular staining on different, alternative sections. This further complicates the visual analysis and integration of molecular and morphological information.

To overcome these problems, we have developed a three-dimensional microscopy system that integrates computer analysis and visualization tools. These tools automate or greatly reduce the amount of interaction required for the acquisition, reconstruction, and morphologically directed analysis of thick tissue samples. Our system can be used to reconstruct tissue structures and to quantitatively measure the presence and spatial distribution of different molecular elements (e.g., genes, RNAs, proteins) in their intact cellular environment. This tool is currently being used to study breast cancer, where heterogeneity and three-dimensionality are at the very base of both disease initiation and clonal progression.

Our system encapsulates a three-dimensional visualization system and an image analysis system. The application was developed using a distributed architecture (client-server model), and Java for writing the graphical user interface (GUI) so that it can run remotely on any computer platform. The system allows

Grant sponsor: Department of Defense Breast Cancer Research Program; Grant numbers: DAMD17-00-1-0306, DAMD17-00-1-0227.

*Correspondence to: Carlos Ortiz de Solorzano, Ph.D., Ms 84-171, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720. E-mail: CODeSolorzano@lbl.gov

Received 20 December 2001; accepted in revised form 27 June 2002

DOI 10.1002/jemt.10233

Published online in Wiley InterScience (www.interscience.wiley.com).

acquiring and registering low magnification (e.g., 1 pixel = 5 μm) conventional (bright field or fluorescence) images of entire tissue sections. It can also create a virtual 3D reconstruction of the tissue structures, from which new areas of interest can be revisited or reacquired at high resolution (e.g., 1 pixel = 0.5 μm) and automatically analyzed.

Although there are existing commercial and noncommercial software packages that can separately perform some of those functions, the integration of all of them (multiresolution, multicolor acquisition of multifield fluorescence microscopy images; reconstruction of structures on interest in 3D; molecular and morphological 3D analyses; content-based image and data storage and retrieval system through the use of "Cases"; or a series of consecutive images and data belonging to a given tissue) on a single platform makes our system a very powerful tool.

A comparison to similar systems shows that many of them offer a set of independent programs running under different interfaces (*IMOD*, Boulder Laboratory for 3-Dimensional Fine Structure), but not a common platform that integrates all of them. Most of these systems are not designed to run on multiplatform environments (VIDA, University of Iowa; Trace, Boston University), some others have special hardware requirements for real time rendering (VoxBlast, Vaytek Inc.) and many lack a proper image management system (Imaris, Bitplane AG). In this review, we describe our system and we illustrate its use by presenting the reconstruction of a mouse mammary gland and a tumor biopsy of a patient with ductal carcinoma in situ (DCIS) of the breast.

## MATERIALS AND METHODS

### System Description

The system is controlled by a client-server application, as shown in Figure 1. The server is a C language application that runs on a computer (Dell Inspiron, running Solaris 7 for Intel) connected to an Axioplan (Zeiss Inc., Germany) microscope. The server actuates all the moving parts of the microscope: motorized scanning stage, excitation filter wheel, and arc-lamp blocking shutter (Ludl Electronic Products Ltd., Hawthorne, NY). It controls the CCD MicroImager camera (Xillix Techonologies Corp., Richmond, British Columbia, Canada) as well. The server can perform basic operations, such as acquiring and storing images, setting the exposure time of the CCD, moving the stage, and operating the filter wheel. In addition, it has been programmed to offer more complex functions, such as automatically focusing the microscope or acquiring multiple field-of-view images. To do this, the server receives each order and divides it into a set of simple actions. For instance, to acquire a multiple field-of-view image, the server asks for the coordinates of the vertices of the area to be acquired and then automatically performs the required sequence of stage movements and camera acquisitions. The output is a mosaic-like image of the area. The server can do multi-color acquisition in fluorescence and bright field, by performing consecutive acquisition using different excitation filters and multi-band emission filters.

***Description of the Client Application.*** The client (R3D2) is connected to the server through UNIX sockets, which are the standard for Internet-based communications. It can send requests to the server from any computer connected to the Internet. To obtain the maximum benefit from this, R3D2 has been written in JAVA (v.1.2), so that it can be executed on many different computer platforms.

Figure 2 shows the R3D2's complete Graphical User Interface (GUI). The interface is divided in two distinguishable parts. One (rightmost vertical panel in Fig. 2) provides connections to the server and allows the user to request its services through a user-friendly interface. The available actions can be classified in two groups.

*Basic control operations.* These include all the simple, atomic, operations provided by the server, such as setting the objective lens, changing the excitation filter (fluorescence), setting the exposure time of the camera, moving the stage to the absolute origin of coordinates relative to which all measurements are taken, opening/closing the arc lamp blocking shutter, and acquiring a single image using the current microscope settings. R3D2 receives the image from the server and displays it, both complete (zoomed out), as well as partially (in its original resolution). Only a small part of the full-resolution image can be displayed at full resolution. The zoomed area can be interactively selected by moving a window on the complete version of the image (Fig. 3). Images can be *saved* both in ICS (Dean et al., 1990) and JPEG format. When images are saved as ICS, all the acquisition parameters (objective, filter, location of the image on the slide are, etc.) are stored in the ICS header file. JPEG format, compressed or not, can be used as an alternative format when the user does not plan any future analysis of the images, and images are stored for exchange, document creation, or web publishing.

*Complex operations.* These operations combine multiple atomic operations to provide the following functionality.

- Autofocus: Automatically focuses the microscope by taking a series of images at different positions in the Z axis (step size = 0.50 μm for low magnification images, 0.25 μ for high magnification images) and determining the best-focused image of the series. Blur, due to out-of-focus light, reduces image contrast, which can be detected using several functions. Based on several comparisons described in the literature (Firestone et al., 1991; Groen et al., 1985; Santos et al., 1997), we selected an autocorrelation-based function introduced by Vollath (1987).
- Scan: Acquires multiple field of view images. The system displays a dialogue-panel where the user can specify the filter(s) to be used (in fluorescence microscopy), exposure time(s) and the limits of the area to acquire. The limits can be defined by its coordinates (when known) or manually, by moving the microscope to the upper, lower, rightmost, and leftmost points of the area.
- Revisit Point: When the user clicks on a point of the image of a previously acquired complete or partial tissue section, the server moves the stage to that location on the slide and takes an image using the current values of objective, filter, and exposure time.
- Revisit Area: When using this option, the user is

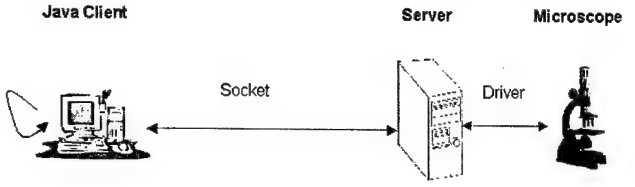



Fig. 1. Description of the client-server architecture of R3D2. The server runs on a computer with a microscope attached, and provides access to the microscope functions. The client is a JAVA application, which can run on any computer (no particular OS required) connected to the server through the Internet Protocol (IP). The communication between client and server uses UNIX sockets. The client provides user-friendly access to all the microscope functions offered by the server, and allows handling of sets of related images (Cases) for storage, annotation, and 3D reconstruction of structures of interest. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

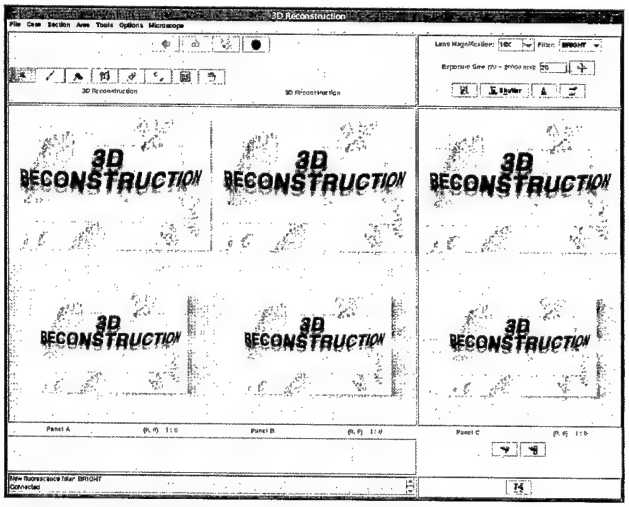



Fig. 2. R3D2's Graphical User Interface. The GUI is divided in two main parts. **Left:** Display consecutive sections of a Case. It also provides the user with tools for registering sections, annotating the images, connecting structures between sections (e.g. mammary ducts, tumor volumes), and reconstructing the annotated images in 3D. **Right:** Access to the microscope related functions offered by the Server. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

asked to draw a rectangle on one image of a previously acquired section. The selected area will be then acquired with the microscope settings provided by the user. Multicolor area acquisition is an option as it is for scanning complete sections.

The second part of the interface (two left vertical panels in Fig. 2) expands the system capabilities, by allowing creating and handling sets of related images, which we call Cases. A Case is a sequence of low-magnification images of complete tissue sections taken from a tissue block, along with all the areas re-visited on them at higher resolution and with different filters, plus the results of analysis performed on them if any. The image files that make up a Case are specially labeled for convenience. The user can:

- Annotate the Cases, by marking and/or delineating structures of interest and linking them within and between consecutive sections. The user can add textual annotations (Text), ductal structure identifiers (with a unique number that identifies them within the section, Duct) and forms that delineate irregularly shaped ducts or other structures (Shapes). In addition to this, ductal marks can be connected within the same section or in different sections (Connected Ducts), and corresponding shapes in different sections can be grouped (Groups).
- Register acquired sections. Before reconstructing a Case in 3D, all its sections must be registered to ensure proper alignment of the elements that will be later reconstructed. For that, we calculate the Rigid-Body Transform that provides the optimum rotation and translation between each pair of sections. The Transform is calculated from three pairs of points interactively marked by the user on each pair of images to be registered. Once the points have been marked, the software calculates the rotation and translation ($\theta$, $t_x$, $t_y$) needed to minimize the sum of the squared distances between all three pairs of corresponding points, thus aligning both images. The results are stored in the second image. This method is very accurate when the pairs of points are spread all along the sections. Reasons for small errors are imprecise mouse interaction, stretching and/or compression of the tissue, and the fact that some structures used to select pairs of corresponding points might not be perpendicular to the sections.
- Reconstruct Cases. Our system reconstructs the tissue structures by rendering the user annotations in 3D. Besides the obvious advantage of volumetric tissue structure visualization, the 3D rendering is linked to the microscope for revisiting, and to the original images and their analysis for display of the images and analysis results, as will be described later. The 3D reconstruction part of the software has been developed using Java3D (v.1.2.1.) This is an application-programming interface (API) for 3D in Java. After asking the user for the range of sections to render, the system converts the coordinates of all the markings in that range of sections from two-dimensional to three-dimensional values. The section number and thickness, along with the distance between sections, determines the depth-coordinate. Then a 3D scene is built using several geometric shapes to represent the different markings. Duct markings are rendered as spheres and connected with lines within and between sections. Contour Shapes are rendered as volumes after applying a refined Delaunay triangulation, using the Nuages reconstruction software (INRIA, France, http://www-sop.inria.fr/prisme/) (Boissonnat et Geiger, 1993). The 3D rendition of the Case is displayed in a new window where the mouse can rotate the scene, zoom in or out or translate the scene in the X and/or Y directions. The 3D window includes a tool bar with options to select elements. By just clicking on one element of the 3D scene, the user can get information about it (location, size, etc.), load the image(s) that contain that element, images are displayed in a new image panel, or move the microscope to its location

on the slide for re-imaging. Volume selection is handled by JAVA 3D. The selection is performed by tracing a "virtual ray" from the user's point of view (defined when rendering the Case, normally at a point corresponding approximately to the position of the user's eye) and the point where the user clicks on the screen. The selected volume will be the first object intersected by the ray within the 3D scene. The user can also hide or show all the different elements of the scene, reset it to the default view and change the scale in any of the three dimensions.

- Analyze Cases. All areas selected based on the 3D morphological reconstruction, can be batch processed upon a user request. This way, only those areas selected based on a particular morphological feature are analyzed, and not all the tissue sections, thus reducing the amount of work required. The analysis is done by streamlining the selected images to a new process running custom-made image analysis routines built on a commercial image processing software (Scilimage, TNO, The Netherlands). At the moment, the image analysis routines can segment counterstained nuclei and detect and quantify FISH probes or punctuate-patterned expressed proteins. The image analysis algorithms for nuclei and signal segmentation have been described elsewhere (Malpica et al., 1997, Ortiz de Solorzano et al., 1998, Malpica et al., 1999). The results of the analysis can be displayed from the 3D rendering window, and global measurements can be performed after selecting the volumes.

An important feature of R3D2 is that all Case-handling and marking functions can be used in parallel to the functions that request microscope actions. Therefore, acquiring a new section can be done in parallel to any other Case related function (e.g., registering already acquired sections or annotating the images). Our implementation of this feature uses Solaris threads. Threads permit executing multiple parallel copies of a program without multiplying resource use. Each thread shares memory and other resources with other threads. R3D2 runs on a main thread and when a microscope-based operation is selected, it launches a new thread that runs on the same memory space as the main one. This scheme guarantees that, in the case of a microscope failure or a socket error, the system will not die abruptly, as only the thread working on the microscope will be affected.

## RESULTS

We will now illustrate the use of our system by showing how two different tissue blocks were imaged and reconstructed. The first one (HB) is a tissue block from the mammary gland of a patient with ductal carcinoma in situ of the breast (DCIS). The second block (MB) is a normal mammary gland of a nuliparus mouse.

### Tissue Source

***HB.*** The tissue was part of a breast lumpectomy specimen. After surgery, the specimen was fixed in an alcoholic-formalin solution, and embedded in paraffin per routine at the California Pacific Medical Center in 1981. The tissue was originally staged as a T1N0M0

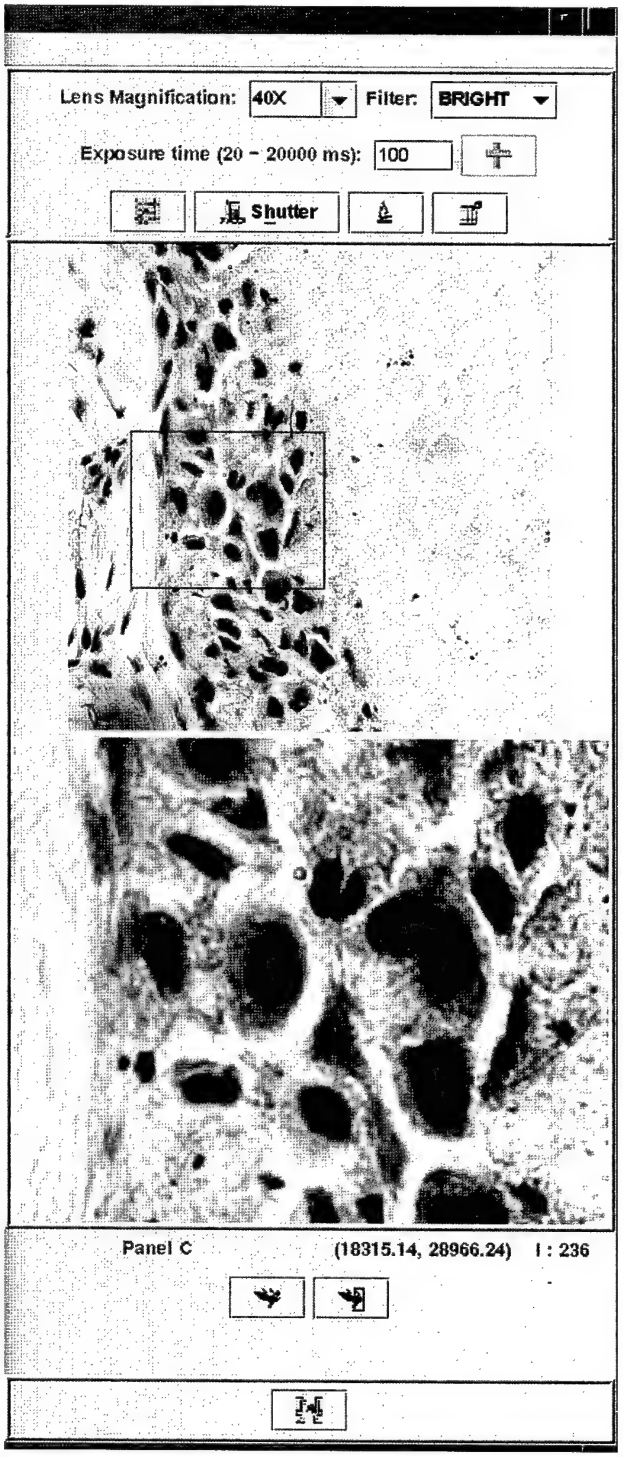



Fig. 3. Single image acquisition. To acquire a single image, the user clicks on the "Acquire" button, after choosing the appropriate microscope settings: objective lens, excitation filter (in fluorescence), and exposure time of the CCD camera. The "Autofocus" option automatically finds the correct focus plane of the microscope before acquiring an image. **Top:** The entire image (reduced to fit in the window). **Bottom:** Contains a zoomed version of the part of the original image selected by the rectangle on the top window. The rectangle can be manually moved to look at different parts of the image at its original, full, resolution. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

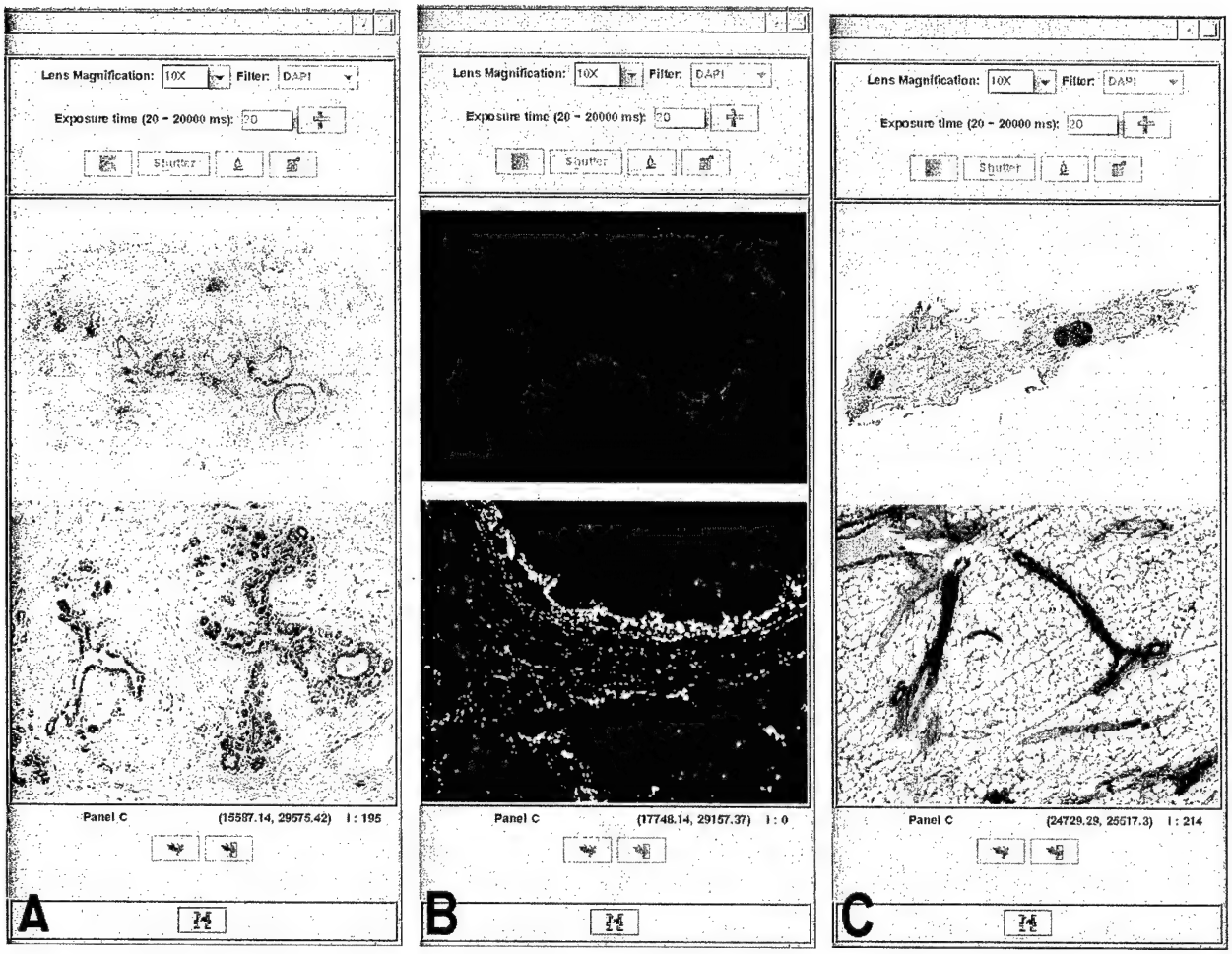



Fig. 4. Tissue acquisition. Images of full sections of a tissue block of ductal carcinoma in situ of the breast (**A,B**) and of a mouse mammary gland (**C**). Both blocks were paraffin embedded, sectioned, and stained for Histology (H&E, images A and C) or counterstained with DAPI (B). Microscope focusing and image acquisition is completely automated, as described in the text. In the individual images, the**top** contains the entire section, the **bottom** contains a zoomed sub-area of it.

(Stage 1). Several blocks of approximately 1" × 1" × 3 mm were taken from the tissue. The block we used contained DCIS and benign ducts; there was no invasive tumor present. The tumor was positive for ERBB2 by immunohistochemistry performed on a 4-μm section taken from the top of the block.

***MB.*** Female BALB/c mice were obtained from Simonson (Gilroy, CA) and housed 4 per cage with chow and water ad libitum in a temperature- and light-controlled facility. Carbon dioxide inhalation was used to kill the animals in accordance with the Association for Assessment and Accreditation of Laboratory Animal Care guidelines and institutional review and approval. The 4th inguinal glands were removed for histology and whole mounts at 10 weeks. The tissue was formalin-fixed and paraffin-embedded.

### Tissue Preparation

***HB.*** The human block was sectioned at 4 μm thickness. Every 5th section was collected onto a plus (treated) slide, for a total of 66 sections. From this set, every other section was collected and stained with H&E (33 sections, 40 μm apart). The rest of collected sections were sent to the UCSF Cytogenetics core for FISH (33 sections, 40 μm apart), with a probe for the ERBB2 gene, which is amplified in 30% of carcinomas of the breast. The ERBB2 probe was RMC17P077, a P1 probe. The DNA was extracted and labeled by Nick translation with red CY3 dUTP fluorochrome. In summary, the odd-numbered sections were H&E stained and the even-numbered sections were DAPI counterstained and processed for FISH.

***MB.*** The mouse gland was paraffin-embedded, H&E stained, and sectioned at 4 μm. We starting collecting every fifth section (20 μm apart); 80 μm into the tissue, we switched and collected all remaining sections (4 μm apart), for a total of 24 sections. The collected sections were all placed on glass slides for microscopy and imaging.

### Imaging of Tissue Sections

***HB.*** Full tissue sections were imaged in bright field (odd-numbered sections) (Fig. 4A), or fluorescence (even-numbered sections) (Fig. 4B), using R3D2's Scan

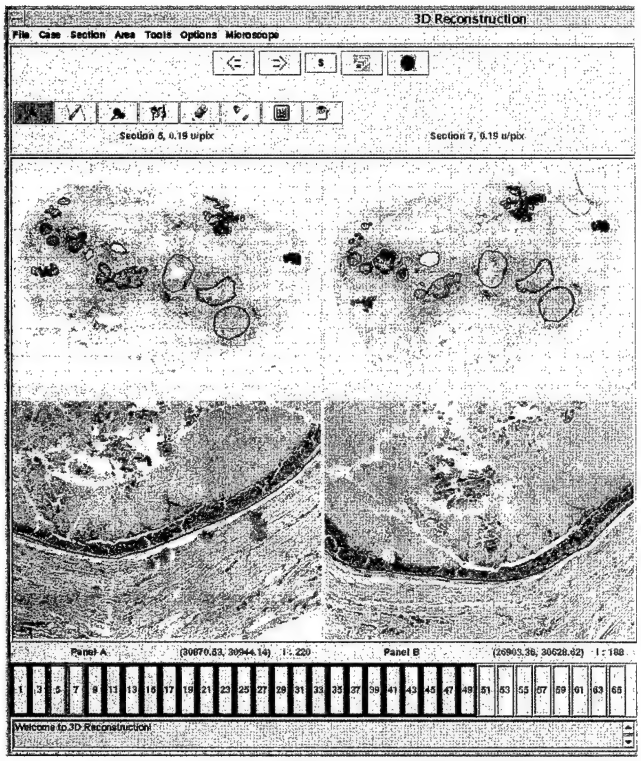



Fig. 5. Tissue annotation (human tissue). Two consecutive H&E stained sections of a human DCIS Case. Only H&E sections are now being shown. The images show manual annotations, namely ducts and tumor masses. Ducts are connected within and between sections. Tumor contours are drawn and also connected between consecutive sections. Connections between sections are shown by changing line color of all connected components every time one of them is visited (selected or just traveled over by the mouse pointer). [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

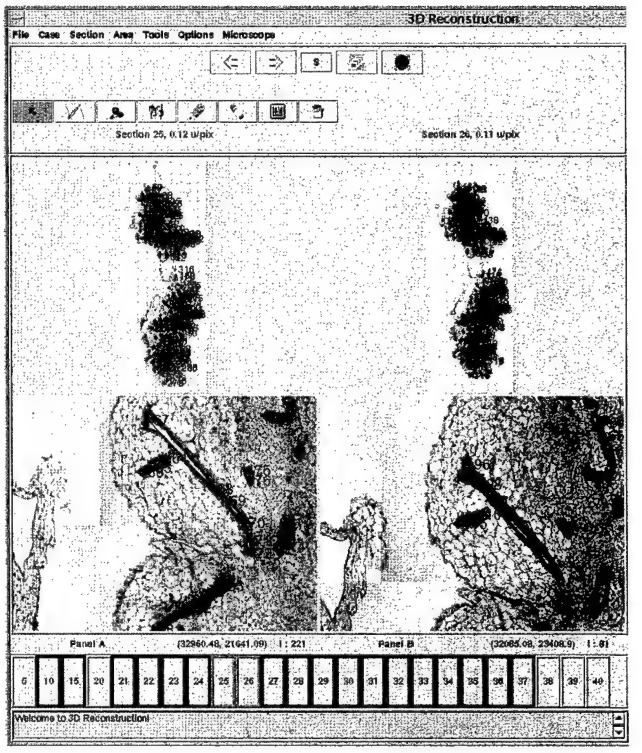



Fig. 6. Tissue annotation (mouse tissue). Two consecutive H&E stained sections of the mouse mammary gland used to test the system. Ducts are connected within and between sections. Lymph node countours are drawn and also connected between consecutive sections. Connections between sections are shown by changing line color of all connected components every time one of them is visited (selected or just traveled over by the mouse pointer). [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

option. Fluorescent sections were imaged using a single-band filter block (360 nm excitation) for DAPI.

***MB.*** All sections were imaged in bright field (Fig. 4C).

In both blocks, sections were imaged at 10× with a Fluar (0.5 n.a.) objective lens (Zeiss, Wetzlar, Germany). In order to optimize memory use while keeping enough resolution for histology, images were reduced by a factor of 4 in both X and Y directions, which gave us an effective 2.5× magnification; i.e., a sixteen-fold reduction in image size. After image compression, the average image size of the sections was 25 Mbytes in the human block and 10 Mbytes in the mouse. The compression was necessary to speed up image transmission and display.

## Creation of Cases

Two Cases were created, initially empty. The acquired sections were added with a number equal to its section number within the tissue. All sections were manually registered as previously described to ensure proper alignment of the to-be-done annotations. Each section was registered to its previous section, thus aligning all sections with the first H&E section of the block.

## Tissue Annotation

***HB.*** All H&E sections were manually annotated. We marked the centers of the ducts by placing a circular mark (R3D2's Duct tool) in the lumen, when the duct was perpendicular to the image plane, or with a line, when the duct was sectioned longitudinally. Then we connected the markings within and between consecutive sections using the Connect Ducts option. We also delineated tumor areas using R3D2's Shapes option and grouped their consecutive sections using the Group option. (Fig. 5).

***MB.*** The same procedure was followed for the mouse except for the Shapes, which were not used to delineate tumor areas, but the lymph nodes (Fig. 6).

## 3D Reconstruction

The reconstruction of the Cases, based on the H&E sections, is shown in Figures 7(HB) and 8 (MB). All the user markings are interactive, in that by clicking on them the user can (1) obtain positional information, (2) load the part of the original image section(s) containing that marking, or (3) revisit the selected marking under the microscope. For the latter, the microscope automatically moves to the position of the marking in the tissue, provided that the right slide is on the microscope. In situ tumors and lymph nodes were rendered as volumes in 3D, and ducts as lines.

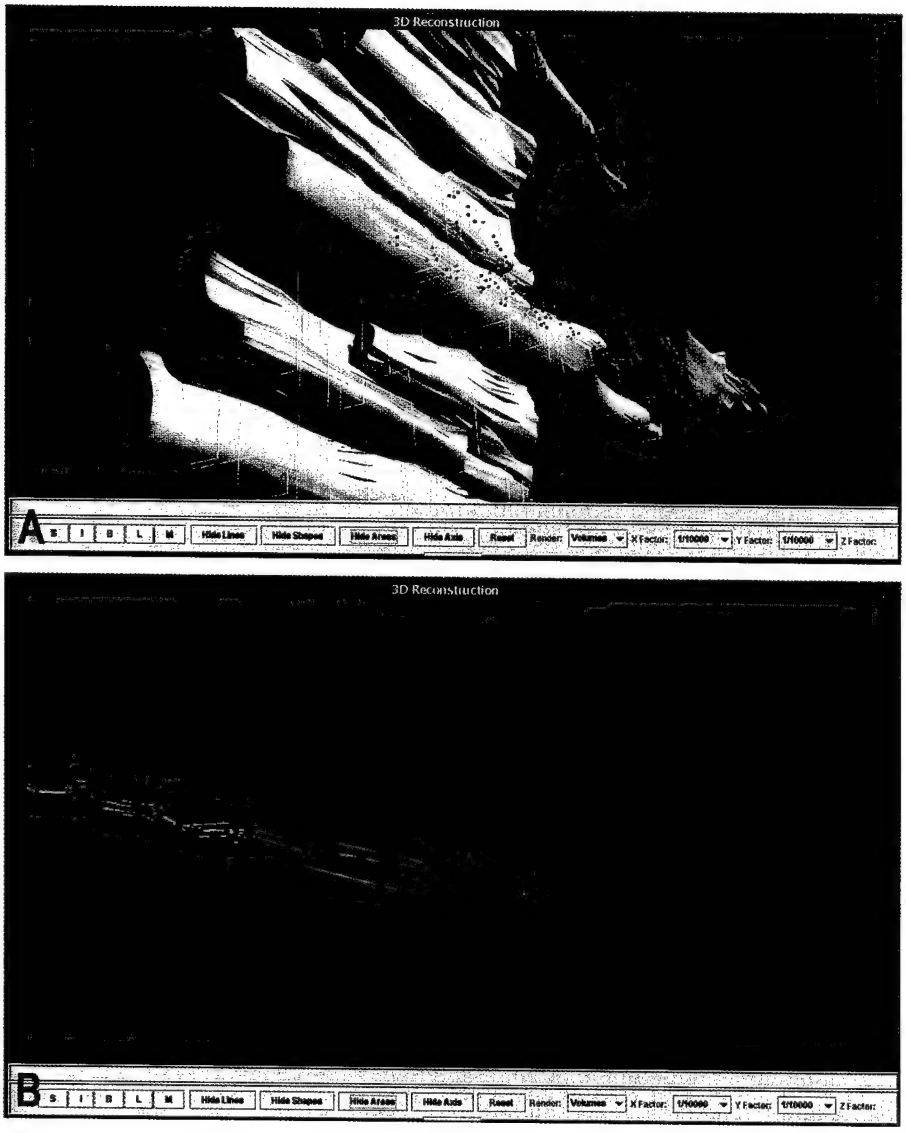



Fig. 7. Tissue reconstruction (human case). 3D reconstruction of the manually annotated DCIS Case. **A:** Tumor volumes have been surface-rendered to show their three-dimensional shapes. Ducts are identified by spheres and connected by lines within and between sections. Yellow rectangles identify areas that were re-acquired at higher magnification. **B:** Close up view of a set of connected markings corresponding to the same duct. The 3D reconstruction is fully interactive. It can be rotated, zoomed, and all the elements can be selected to retrieve information about the element, display the original image that contains the element, or automatically move the microscope to the selected point.

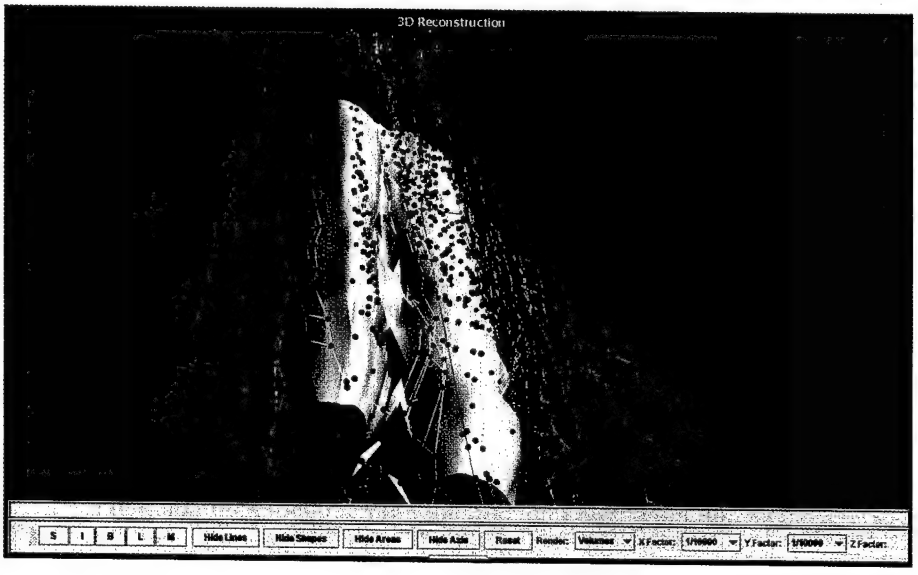



Fig. 8. Tissue reconstruction (mouse Case). 3D reconstruction of the H&E stained, manually annotated mouse Case used to test our system. Lymph node volumes have been surface-rendered to show their three-dimensional shapes. Ducts are identified by spheres and connected by lines within and between sections.

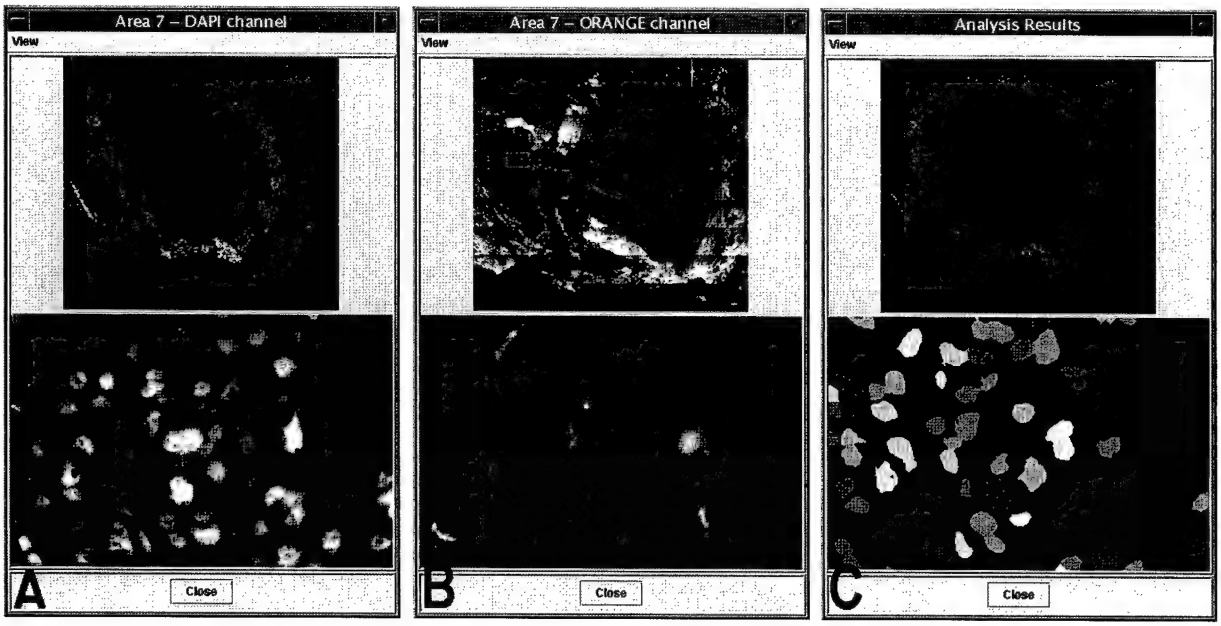



Fig. 9. Revisiting and analysis of areas of interest. Revisited area of the Human DCIS tissue block used to test the system. **A:** DAPI image of counterstained nuclei; **B:** CY3 images of FISH with a probe to the erb-b2 gene. Areas can be acquired and loaded from the images or the sections they belong to or directly from the 3D reconstruction of the tissue. **C:** Results of the segmentation of nuclei based on the counterstained channel. Each segmented nucleus is colored differently. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

### Revisit of Areas of Interest

***HB.*** Several areas were revisited at high-resolution (40×), with a Plan-Neofluar (1.3 na) oil immersion lens (Zeiss, Wetzlar, Germany). Areas were manually selected having either morphologically normal or abnormal (DCIS) areas. Areas were double-scanned using a Pinkel filter set (Chroma Technologies, Brattleboro, VT), by consecutively imaging while exciting the sample at 360 (DAPI emission from the nuclei) and 572 nm (CY3 emission from ERBB2 gene). All areas were acquired at full resolution and then compressed by a factor of 2 in both X and Y directions for an effective 20× magnification. Areas were manually selected by drawing rectangles on the low-resolution images of the whole fluorescent sections, although they could have been selected from the 3D reconstruction as well. To ensure proper alignment between the areas and the sections, the images of the sections were previously aligned with the microscope slide by calculating the shift between the current location of the slide and its location when the image of the section was originally acquired. This is done by calculating the rigid body transform between three points manually selected in the image of the section and the corresponding points under the microscope. Figure 7 shows the 3D reconstruction of the human Case, incorporating the high-resolution acquired areas, displayed as rectangles. As all other elements in the reconstruction model, the areas can be loaded or revisited by clicking on them (Fig. 9).The total number of areas imaged was 160.

### Analysis of Areas of Interest

***HB.*** All 160 areas were segmented overnight using the Analyze Case function. In this particular case, DAPI counterstained nuclei were segmented using a two-step algorithm that first uses an adaptative threshold to separate DNA areas from the background and then applies a Hough-transform + Watershed algorithm to separate clusters of nuclei resulting of the adaptative thresholding. FISH signals were segmented using a TOP-HAT morphological algorithm followed by morphological reconstruction. The FISH segmentation algorithm was applied only to those areas identified as nuclei on the DAPI channel. A detailed description of these methods is out of the scope of this study and can be found in the literature (Malpica et al., 1997, Ortiz de Solorzano et al., 1998, Malpica et al., 1999).

The total number of nuclei segmented was above 200,000. The results of the nuclear and FISH segmentation can be displayed for every area (Fig. 9) from the images of the sections they belong to or directly from the 3D rendering window.

## DISCUSSION

We have presented a new and powerful computer-based system that allows automation of the acquisition, storage, and analysis of thick sectioned tissue specimens. The system has been described and demonstrated by reconstructing tissue blocks from two tissue sources, processed using different staining and microscopy protocols.

By integrating information from different types of staining, both histological (e.g., H&E) and molecular (e.g., IHQ or FISH), we allow simultaneous morphological and molecular analysis of the specimens. The molecular analysis is not only simultaneous, but in fact driven by the morphology as the areas acquired and analyzed can be selected directly from the 3D recon-

struction of the structures marked on the low-resolution images. This way, the labor-intensive task of acquiring and analyzing the images is done semi-automatically—the slides still need to be manually placed on the microscope and the areas to be acquired drawn on the images of the sections—enormously reducing the time and labor. In fact, large-scale analysis can be done (e.g., whole gland analysis), which would be unthinkable otherwise.

A classical problem of fluorescence microscopy, which limits the number of labeled elements to the number of fluorochromes that can be discriminated from, can be overcome by using single or dual color staining on several adjacent sections, provided that the distance between histological (e.g., H&E) sections enclosing them allows detection of continuity between the structures of interest.

Future work will involve adding new functionality to the system by augmenting the number of analysis functions provided (e.g., detection of cytoplasmic or extra cellular proteins), and automating the detection of the structures of interest, which at this point is the most time-consuming task when reconstructing a Case. Interaction could be further reduced by using an automatic slide feeder attached to the microscope, which would eliminate the manual loading of slides for revisiting or acquisition of areas.

## ACKNOWLEDGMENTS

We thank K. Chew and S. Ravani for their help selecting and coordinating tissue sectioning and staining, Drs. G. Maggrane and K. Ewan for their work on the IHQ and FISH staining of the sections, Drs. F. Waldman, K. Chin, and J. Gray for their support and advice regarding the selection of the tissue, and A. Harris for her help with tissue imaging and annotation.

## REFERENCES

Boissonnat JD, Geiger B. 1993. Three dimensional reconstruction of complex shapes based on the *Delaunay* triangulation. In: Acharya RS, Goldgof DB, editors. Biomedical image processing and biomedical visualization. SPIE 1905:964–975.

Dean P, Mascio L, Ow D, Sudar D, Mullikin J. 1990. Proposed standard for image cytometry data files. Cytometry 11:561–569.

Firestone L, Cook K, Culp K, Talsania N, Preston K. 1991. Comparison of autofocus methods for automated microscopy. Cytometry 12:95–206.

Groen F, Young IT, Ligthhart G. 1985. A comparison of different focus functions for use in autofocus algorithms. Cytometry 12:81–91.

Malpica N, Ortiz de Solórzano C, Vaquero JJ, Santos A, Vallcorba I, García-Sagredo JM, del Pozo F. 1997. Applying watershed algorithms to the segmentation of clustered nuclei: defining strategies for nuclei and background marking. Cytometry 28:289–297.

Ortiz de Solórzano C, Santos A, Vallcorba I, García-Sagredo JM, del Pozo F. 1998. Automation of FISH Spot counting in interphase nuclei: statistical evaluation and data correction. Cytometry 31:93–99.

Ortiz de Solórzano C, García Rodriguez E, Jones A, Sudar D, Pinkel D, Gray JW, Lockett SJ. 1999. Segmentation of confocal microscope images of cell nuclei in thick tissue sections. J Microsc 193:212–226.

Santos A, Ortiz de Solorzano C, Vaquero JJ, Pena JM, Malpica N, del Pozo F. 1997. Evaluation of autofocus functions in molecular cytogenetic analysis. J Microsc 188:264–272.

Vollath D. 1987. Automatic focusing by correlative methods. J Microsc 147:249–288.

# Automatic segmentation of histological structures in normal and neoplastic mammary gland tissue sections

Rodrigo Fernandez-Gonzalez *[a, c], Thomas Deschamps[b], Adam Idica[a],
Ravi Malladi[b], Carlos Ortiz de Solorzano[a]
[a]Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720
[b]Mathematics Department, Lawrence Berkeley National Laboratory, Berkeley, CA
[c]UC Berkeley /UC San Francisco Joint Graduate Group in Bioengineering, Berkeley, CA 94720

## ABSTRACT

In this paper we present a scheme for real time segmentation of histological structures in microscopic images of normal and neoplastic mammary gland sections. Paraffin embedded or frozen tissue blocks are sliced, and sections are stained with hematoxylin and eosin (H&E). The sections are then imaged using conventional bright field microscopy. The background of the images is corrected by arithmetic manipulation using a "phantom". Then we use the fast marching method with a speed function that depends on the brightness gradient of the image to obtain a preliminary approximation to the boundaries of the structures of interest within a region of interest (ROI) of the entire section manually selected b the user. We use the result of the fast marching method as the initial condition for the level set motion equation. We run this last method for a few steps and obtain the final result of the segmentation. These results can be connected from section to section to build a three-dimensional reconstruction of the entire tissue block that we are studying.

**Keywords:** 3D reconstruction, background correction, brightness gradient, fast marching method, level set method

## 1. INTRODUCTION

The mature mammary gland has been characterized as a tree-like structure formed by hollow ducts converging at the nipple. From there a primary duct branches into secondary ones, which themselves form tertiary ducts. The ducts, lined by one or two layers of epithelial cells, finish in secretory structures known as terminal ductal lobular units, which are the site of milk production after pregnancy.

In cancer, this hierarchical structure is disrupted by an uncontrolled growth of the ductal epithelium from the lobula units into the ducts and the subsequent invasion of the surrounding tissue. Thus, morphological or structural changes accompany the genetic changes that occur in breast cancer.

An interesting problem is posed by the idea of quantifying these genetic or molecular changes that occur during breast cancer progression in the context of the tissue environment where they occur. This would allow studying the correlation between morphological and molecular mutations in breast cancer. Since the mammary gland and its constituent elements are neither flat nor homogenous, the approach to this problem should be three-dimensional, and take into account the heterogeneity of the gland. However, most biological methods focus on just one type of changes (molecular or morphological) and neglect three-dimensionality and heterogeneity.

Although imaging tissue structure and function *in vivo* would be extremely desirable, the existing *in vivo* imaging methods (X-ray, MRI, optical tomography, etc) don't provide the necessary resolution or penetration into the tissue fo cell-level molecular analysis and/or require an imaging-forming substrate that can not be introduced in the tissue without severely altering its normal functioning. Consequently, *ex-vivo* microscopic analysis of the tissue from flat fixed sections is the routine method in histopathology. However, our limited ability for extrapolating and visualizing three-dimensional information from sequences of two-dimensional scenes renders this method non suitable for quantitative three-dimensional tissue characterization.

To overcome these issues we have developed a system for simultaneous morphological and molecular analyses of thick tissue samples[1]. The system is composed of a compute-assisted microscope and a JAVA-based image display, analysis

* rfgonzalez@lbl.gov

and visualization program that allows semi-automatic acquisition, annotation, storage, three-dimensional reconstruction and analysis of histological structures of interest (intraductal tumors, normal ducts, blood vessels, etc.) in thick tissue specimens. For this the tissue needs to be extracted from the donor and embedded in a permanent or semi-rigid medium before being fully sectioned. Then the sections have to be stained in a way that visually highlights the desired structures.

In this paper we will focus on the annotation of the structures of interest on the H&E stained sections. Manual segmentation has been used before to delineate histological structures[2, 3] and therefore we initially annotated each of the interesting features of the tissue manually on each section in order to build the three -dimensional reconstruction of the block. This constituted the bottleneck in the study of samples. Semiautomatic approaches to segmentation of features of interest in histological sections have also been used[4], but they still involve a lot of user interaction, which remains the bottleneck in the process of analyzing samples. Automatic annotation (segmentation) of the structures of interest is the answer to this problem. Here we propose an automatic method followed by interactive correction that greatly reduces the interaction required, thus eliminating the b ottleneck and opening the door to the use of our system for imaging and reconstruct large, complex tissue structures.

Active contour models[5] have been used before to segment objects from image data. These methods are based on deforming an initial "seed" c urve towards the boundary of the desired object. The deformation is achieved b minimizing a certain energy function that approaches a minimum near the object boundary. However, these models onl produce satisfactory results for smooth objects, failing to correctly delineate objects with big protrusions. Due to the existence of multiple local energy minima for this kind of objects, the result of the segmentation largely depends on ho good the initial condition is, i.e., how close is the initial seed curve to the desired contour. Moreover, active contours are topologically constrained and cannot deal with multiple objects in a single image[6], thus requiring running an independent minimization process for each object/seed pair.

Given the limitations of all th e above-mentioned methods, we proposed to use the level set approach[7, 8] which is an interface propagation technique used for different applications, including segmentation. The initial curve is represented here as the zero level set of a higher dimensiona l function, and the motion of the curve is embedded within the motion of the higher dimensional surface. The speed of that motion is defined based on the characteristics of the image to be segmented, and thus adjusted so that when the interface passes ove low brightness gradient areas the interface expands quickly, whereas when the gradient is large (indicating the location of an edge in the image) the curve is slowed down. In addition, a surface tension term (in the form of curvature) is included in the s peed function to slightly retard the expanding contours at very high -gradient areas, which usually correspond to image noise. The interface can propagate within narrow regions, and is able to handle topological changes, thus being able to segment multiple objects simultaneously. In addition, handling an inherently two-dimensional problem in three dimensions by embedding the seed into a continuous surface facilitates calculating the curvature of the front and dealing with singularities in the evolution of the curve.

However, changing an n-dimensional problem into one in n+1 dimensions increases the computational cost. The narro - band method accelerates the level set flow by updating the position of the curve only in a narrow vicinity of its current location. But in our experience, the narrow band technique does not reduce the computational cost to a reasonable limit, due to the large size of the images of the sections. Thus, we propose to use the fast marching method [9, 10], a numerical technique to solve the equation that drives the movement of the curve by combining an efficient –constrained- solution to the equation of the movement of the front, narro -band level set methods and a *min*-heap data structure. This method is only used for monotonically advancing fronts, providing a very fast solution, even though not as accurate as the one obtained by using the level set method. This solution is then used as the initial condition for the slower but more accurate full level set segmentation.

The rest of this paper is organized as follows: Section 2 describes the general tissue handling and image acquisition that we use, and the theoretical basis of the segmentation scheme. Section 3 shows the results of applying our method to histological tissue sections. Section 4 discusses the results and suggests some future developments and improvements to our technique.

## 2. METHODOLOGY

### 2.1 Tissue processing and imaging

The tissue processing and staining protocol that we use for the analysis is illustrated in figure 1. Tissue b locks of 4 to 5 mm of thickness are sliced into 5 μm (thin) sections. The odd sections are stained with H&E, which is the standard staining in routine histopathology. It provides morphological information both at the cytological (single cell) and architectural (organ) level. The even sections are stained using some fluorescent technique ( *in situ* hybridization, immunocytochemistry, ...) depending on the molecular phenomena that we want to study. Describing the acquisition and analysis of the fluorescent section is out of the scope of this paper. What follows, therefore, describes only the protocol used for the H&E stained sections. Low magnification (2.5X) images of all the sections are automaticall acquired using a motorized Zeiss Axioplan I microscope coupled with a monochrome XilliX Microimager CCD camera. This is done by automatically scanning the area of the slide occupied by tissue, and tiling together all the individual snapshots into a single whole-view image of the sections. The sequence of microscope movements and camera operations is determined by an application running on the UNIX Ultra 10 workstation that drives both the camera and all moving parts of the microscope.

The next step consists of annotating the structures of interest (ducts, lymph nodes, tumors, ...) in the images of the H&E stained sections. These annotated structures are then used to render a three-dimensional model of the tissue block where the morphology of the sample helps in determining which are the areas of the tissue where a mole cular analysis may be more interesting. After selecting these areas in the three-dimensional rendition of the organ, the system asks the user to place the right fluorescent section(-s) on the stage, and high magnification (40-100X) images of the chosen areas are acquired. Finally, quantification routines are run on these new images and results are collected.

Manually annotating tissue structure in complex sections, in tissue blocks containing at least tens of sections, is feasible but for all purposes imp actical due to the tremendous human effort involved. Although it might be the most accurate and reliable approach, it is not possible to carry out except when reconstructing small, simple tissue volumes. In consequence, we have developed automatic methods that eliminate or greatly reduce human interaction, thus making the reconstruction of complex systems possible. What follows describes our approach.

### 2.2 Image Processing

#### 2.2.1 Preprocessing

The algorithm that we use to acquire images of entire sections c eates a mosaic from a set of snapshots, each of which is the image of a single field of view. This approach causes the formation of a background pattern throughout the image (figure 2a), involving relatively large brightness gradients in between elements o f the final mosaic. Now, there is no reason why an object of interest could not expand over several fields of view. Since our segmentation approach depends largely on the brightness gradients of the image, we need to eliminate the background pattern in order to obtain good segmentation results.

We achieve this goal by executing a set of arithmetic operations on the "mosaic" image. First we need a "phantom" image, that is, an image of an empty field of view taken under the same illumination conditions and m icroscope configuration that we used to acquire the initial image. Since all the images that we acquire have an empty frame in the upper left corner, we chose that frame as our "phantom" for the corresponding section. We then compute the average brightness of that "phantom" frame and add it to each pixel of the initial mosaic. We invert each frame of the mosaic and subtract from each one of them the inverse of the phantom image. The resulting mosaic is an inverted, background-corrected image. By doing a final inversion we get the image we were looking for (figure 2b).

#### 2.2.2 Preliminary segmentation: Fast marching method

To obtain a first approximation of the structures that need to be segmented we use the fast marching method. We will now review the basics of this method. Consider a monotonically advancing front with speed $F$ always positive (it could also be always negative) in the normal direction. Let $T(x, y)$ be the time at which the curve crosses the point $(x, y)$. It is obvious that the gradient of the a rival time is inversely proportional to the speed of the function, that is,

$$|\nabla T|F = 1. \quad (1)$$

This is a form the Eikonal equation, which is very closely connected with the notion of viscosity solutions. These solutions can be achieved either by iteration or by stationary construction of $T(x, y)$. The algorithm that we use relies on the later approach. If we assume that our domain is the unit box, and imagine that the initial front is along the line $y = 0$, we can see that the front propagates upward off the initial line. Using the right approximation to the gradient, we are looking for a solution in the domain to:

$$\left[\max\left(D_{i,j}^{-x}T,0\right)^2 + \min\left(D_{i,j}^{+x}T,0\right)^2 + \max\left(D_{i,j}^{-y}T,0\right)^2 + \min\left(D_{i,j}^{+y}T,0\right)^2\right]^{\frac{1}{2}} = \frac{1}{F_{i,j}}, \quad (2)$$

where $D^-$ and $D^+$ are backward and forward difference operators. This is in essence a quadratic equation for the value at each grid point, and thus we can solve the equation at each of these points and take the largest possible value as the solution. Hence, the algorithm rests on building a solution outwards from the smallest $T$ value. The building zone is confined to a narrow band around the front.

First we need to initialize three lists: *Visited* contains the points $(i, j)$ at which the value for $T$ is already fixed (initiall just the seed points with $T = 0$); *Close* contains the set of all grid points in the narrow band, whose $T$ value we set as inversely proportional to $F_{i,j}$; and *Far* contains the rest of the grid points ($T = \infty$). After this we start marching forward. We remove the point in *Close* with smallest $T$ value and we add it to *Visited*. Then we move all (if any) of the 4-connected neighbors of this point that are in *Far* to *Close*, and finally recomputed the $T$ value for all the points in *Close* before we iterate (look for the point in *Close* with the smallest $T$ value, ...). Figure 3 shows a snapshot of a 2D grid after some iterations of this algorithm.

It is evident that an efficient version of this algorithm relies on being able to find the point with smallest value fo $T$ in the narrow band as fast as possible. Using a *min*-heap data structure to store the value s of $T$ for the points in the narrow band together with the indices that give the location of those points on the grid we achieve that goal. The *min*-heap is a complete binary tree where the value at any given node is less than or equal to the values at its children. Therefore, the point with smallest $T$ value is always located at the root of the tree.

Finally, we define the speed of propagation in the normal direction as a decreasing function of the brightness gradient $|\nabla I(\mathbf{x})|$, that is, a function that is very small near big image gradients (i.e. possible edges) and vice versa:

$$F(x) = e^{-\alpha|\nabla I(\mathbf{x})|}, \alpha > 0, \quad (3)$$

where $\alpha$ is the "edge strength", or importance that we give to the presence of a brightness gradient in order to slow down the front. Depending on this value the speed function falls to zero more or less rapidly, and thus, it could stop few grid points away from the real edge. Also, variations in the gradient along the boundary can cause inaccurate results. False gradients due to noise can be avoided using an edge preserving smoothing scheme on the image as a preprocessing step.

The user can run the fast marching flow from a given set of initial seeds (mouse clicks on the background of the images). Alternatively, the user can decide to segment only the structures within a manually defined rectangular region of interest. If the region is too big, subsampling can be used so that the segmentation process is not to slow. This option must be used carefully, since subsampling distorts the definition or completel obliterates small structures. The resulting contou (or contours in the case that we are segmenting several objects at the same time) will provide an excellent initial condition for the level set method.

Putting al these elements together we can get a goo d approximation of the shape of the object that we are trying to segment (figure 4). In order to improve the final segmentation we propose to run a few iterations of the level set method using the result of the fast marching method as the initial condition .

### 2.2.3 Final segmentation: Level set method

Once we have obtained a good approximation of the shape of the object using the fast marching method, we can afford to use the more computationally expensive level set method to improve the result of the segmentation. The essential idea here is to embed our marching front as the zero level set of a higher dimensional function $\psi(x, y, z) = \pm d$, where $d$ is the signed distance from $(x, y)$ to the front (figure 5). Using the chain rule we can find the motion equation:

$$\psi_t + F(x, y, z)|\nabla\psi| = 0, \quad (4)$$

where $F$ is again the speed in the normal direction and the subscript denotes a partial derivative with respect to time. This is an initial value partial differential equation, since it describes the evolution of the solution on the basis of an initial state. As pointed out before, the level set approach offers several advantages: the zero level set of the function can change topology and form sharp corners; a discrete grid can be used together with finite differences to approximate the solution; intrinsic geometric quantities like normal and curvature can be easily extracted from the higher dimensional function; and everything extends directly to three dimensions.

If we now let the surface move with a speed $F = 1 - \varepsilon K$, where $K(\mathbf{x})$ is the curvature and $\varepsilon > 0$, we have a two component speed function. The constant term of $F$ causes the front to seek object boundaries, and the curvature component controls the regularity of the deforming shape: it forces relaxation of the contour and smoothing of sharp corners. The bigger we make $\varepsilon$, the more importance we give to this "motion by curvature" term.

To mold the initial contour (in our case the result of the fast marching method) into the desired shape we use two force terms. One of them is very similar to the one shown in equation 3, and the other one attracts the front towards the object boundaries, providing an stabilizing effect when the brightness gradient has large variations. By substituting these two terms in the motion equation we get:

$$\psi_t + g_I(1 - \varepsilon K)|\nabla\psi| - \beta\nabla P \cdot \nabla\psi = 0, \quad (5)$$

where,

$$g_I(\mathbf{x}) = \frac{1}{1 + |\nabla I(\mathbf{x})|}, \; P(\mathbf{x}) = -|\nabla I(\mathbf{x})|. \quad (6)$$

As before, $I(\mathbf{x})$ is the image, optionally smoothed using an edge-preserving smoothing scheme. $P$ is a potential field that attracts the front towards edges in the image. This attraction is controlled by a parameter, $\beta$. $\nabla P \cdot \nabla\psi$ is the projection of an attractive force vector on the surface normal.

We can use the result of the fast marching method, $T(x, y)$, as the initial condition for the level set algorithm: $\psi(x, y; t = 0) = T(x, y)$. By solving equation 5 for a few time steps using the narrow band approach as well, we obtain an accurate segmentation of the desired object (figure 6).

## 3. RESULTS

In this section we consider the problem of reconstructing some intaductal cancer areas in a tissue biopsy of a cancerous human breast and a set of ducts in an entire mouse mammary gland. The tissue samples were sliced, and we used 56 sections for the human case and 70 for the mouse one. Therefore, manually segmenting each of the structures in every section is an extremely expensive process. To automatically segment them using the fast marching-level set method previously described, we begin by defining a ROI where we will run the segmentation. This ROI can be extended to cover the entire section, but then, considering the size of the images, it is wise to use subsampling in order to run the algorithm in real time.

Then we are presented with a panel (figure 7) where we can select the parameters of the algorithm. We have observed that a particular set of parameters is frequently good enough to segment similar structures (i.e., all the tumors, or all the ducts, or all the lymph nodes, ...) throughout all the sections of a particular tissue block. The set of parameters that can be customized are:

- *Delta t*: how often we update the level set function inside the narrow band.
- *Move fronts outwards/inwards*: direction of expansion of the fronts, i.e., sign of the speed function. Inwards expansion is used when the initial seed is a polygon drawn around the structure of interest. Otherwise, outwards is used most of the time.
- *Epsilon*: weight of the "motion by curvature" term in the motion equation of the level set algorithm. The bigge this value, the more relaxation we force on the expanding front, and the smoother the result.
- *Beta*: weight of the attraction that edges exert on the fronts in the level set motion equation.
- *Narro -band width*: normalized width of the narrow band used in both the fast marching method and the level set algorithm. Higher values render more accurate results, at the expense of increased computations.
- *FMM time*: how long the fast marching method will run.
- *Number of iterations of LS method*: how many iterations of the level set method to run. Normally, the number of steps positively correlates with the accuracy of the final segmentation. However, an excess in the number of steps, besides dramatically increasing computing time, can produce oversegmentation if the convergence of the algorithm is not guaranteed due to incomplete contours or image artifacts
- *Min. contour length*: size filter: size (in pixels) of the smallest segmented object that will not be discarded as noise.
- *Max. contour length*: size filter. This is the maximum size allowed for a segmented object.
- *Use FMM*: determines if the fast marching method should be used to obtain a preliminary result or not. At times, for small features, it is possible to just use the level set method from the initial seed.
- *Subsample*: determines whether the ROI must be subsampled.
- *Subsampling level*: amount of subsampling. More should be used with increasing image sizes, without compromising the resolution and level of detail of the final segmentation. An option would be to use an "imperfect" result, calculated from the subsampled image as the initial condition for a few steps of the level set method on the full resolution image.
- *Preprocessing*: The user can select to de-noise the original image using Gaussian smoothing.

After selecting the parameters of the flow, seed points are defined inside (to find the internal contour) or outside (to find the external contour) of the structures of interest. In most cases one click is enough. Large images may require several, evenly distributed clicks. The value of $T(\mathbf{x})$ at these points is set to zero, and the initial heap is constructed from thei neighbors in order to start running the fast marching method. When this method finishes, the final $T(\mathbf{x})$ function is passed as the initial condition to the level set motion equation (eq. 5). We solve that equation for a few steps. This segmentation scheme provides a result in less than one second for images whose size (after subsampling, if any) is around 2 kilobytes (e.g., 512 x 512 pixels) running on a Sun Ultra 10 workstation with 1 GB of RAM.

Figures 8 and 9 show two examples of the results obtained with the combination of the fast marching and the level set methods. Figure 8b shows the segmentation of a tumor mass in a human tissue block. The results of the segmentation can be edited and removed with the interactive tools provided by our system (figure 8c). An area was selected around the structures of interest in the section. No subsampling was used. The initial seeds are represented by blue points in figure 8a. In figure 9 we can see the segmentation of the external contours of three ducts in a particular area of an entire mouse mammary gland. In this case we also run the algorithm on a ROI on one of the sections with no subsampling factor. Figure 9a shows the points that were used as initial seeds. Figures 9b and 9c show the results of the segmentation before and after interactive correction.

Finally the segmented shapes are connected (manually) between sections, and three-dimensional reconstructions of the samples are built. Figure 10 shows a reconstruction of the tumors contained in the human tissue block that contains the tumor shown in figure 8. Surface noise can be reduced by increasing the "motion by curvature" term, as described in the methods section. In figure 11 we can see the reconstruction of the normal ducts segmented in the images of the mouse mammary gland (see figure 9).

## 4. DISCUSSION

Three-dimensional reconstruction of thick tissue-blocks from serial sections is an important aid for morphological analysis of normal and diseased mammary gland specimens. In this paper we have presented a method that reduces the time and interaction needed to build the three -dimensional model of a tissue block, enormously increasing the throughput of the analysis and by doing so making it feasible to use this a pproach for the analysis of complex, large specimens. To achieve this goal we have combined image processing techniques and two well -established schemes fo interface propagation: the fast marching method and the level set method.

Our approach starts by correcting the background of the images. This is an important step, since the background pattern generated when an image is acquired modifies the brightness gradient of the image, and the speed functions that we use for interface propagation depend on that gradient. Once the background has been corrected, we run the fast marching method. This technique provides a good approximation of the boundaries of the objects that we are trying to segment in a very short time, since it assumes monotonic speed functions (always positive or always negative). We then use the approximation provided by the fast marching method as the initial condition for the level set method. This more computationally expensive algorithm is run just for a few steps, which are enough to fit t he front to the contours of the structures of interest, but not as many as to make the segmentation too expensive. Finally, the option to subsample large images can enormously reduce the time required for the segmentation.

This approach can still be improved. We are currently working on a time-step independent scheme that is expected to be faster than the current one. To improve the accuracy of the results, we have developed an edge -preserving smoothing algorithm based on Beltrami's flo [11], which will replace the Gaussian smoothing currently used before executing the segmentation methods. This algorithm eliminates false gradients due to noise while enhancing gradients due to object boundaries, thus allowing the front to fit the boundaries of the object more accurately. Finally, the level set approach is readily extensible to 3D. Being able to segment three-dimensional structures of interest versus two-dimensional ones would save the process of connecting the segmented 2D contours from section to section, thus improving the analysis time. Also, since geometric properties can be easily extracted from the higher dimensional function, we could easil obtain some information about the extracted volume from the segmentation algorithm itself.

## ACKNOWLEDGEMENTS

The U.S. Army Medical Research Materiel Command under grants DAMD17-00-1-0306 and DAMD17-00-1-0227 and the Lawrence Berkeley National Lanoratory LDRD program supported this work.

## REFERENCES

1. Fernandez-Gonzalez R., Jones A., Garcia-Rodriguez E., Chen P.Y., Idica A., Barcellos-Hoff M.H., Ortiz de Solorzano C. "A system for combined three-dimensional morphological and molecular analysis of thick tissue samples", *Microscopy Research and Technique*, Vol. 59(6), pp. 522-530, December 2002.

2. Moffat D. F., Going J. J. "Three dimensional anatomy of complete duct systems in human breast: pathological and developmental implications", *Journal of Clinical Pathology*, 49, pp. 48-52, 1996.

3. Ohtake T., Kimijima I., Fukushima T., Yasuda M., Sekikawa K., Takenoshita S., Abe R., "Compute -assisted complete three-dimensional reconstruction of the mammary gland ductal/lobular systems", *Cancer*, Vol. 91(12), June 15, 2001.

4. Manconi F., Markham R., Cox G., Kable E., Fraser I. S. "Compute -generated, three-dimensional reconstruction of histological parallel serial sections displaying microvascular and glandular structures in human endometrium", *Micron*, 32, pp. 449-453, 2001.

5. Kass M., Witkin A., Terzopoulos D. "Snakes: Active contour models". *International Journal of Computer Vision*, Vol. 1, pp. 321-331, 1988.

6. Malladi R., Sethian J. A., Vemuri B. C. "Shape modeling with front propagation: a level set approach". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 17(2), pp. 158-175, Feb. 1995.

7. Osher S., Sethian J. A. "Fronts propagating with curvature dependent speed: algorithms based on Hamilton-Jacobi formulation". *Journal of Computational Physics*, Vol. 79, pp. 12-49, 1988.

8. Sethian J. A. "Level Set Methods: An Act of Violence", *American Scientist*, Vol. 85 (3), May-June 1997.

9. Sethian J. A., "A marching level set method for monotonically advancing fronts", *Proceedings of the National Academy of Sciences*, Vol. 93(4), 1996.

10. Malladi R., Sethian J. A. "A real time algorithm for medical shape recovery", *Proceedings of International Conference on Computer Vision,* pp. 304-310, Mumbai, India, January 1998.

11. Kimmel R., Malladi R., Sochen N. "Image Processing via the Beltrami Operator", *Proceedings of Asian Conference on Computer Vision,* Vol. 1351, pp. 574-581, Hong Kong, January 1998.

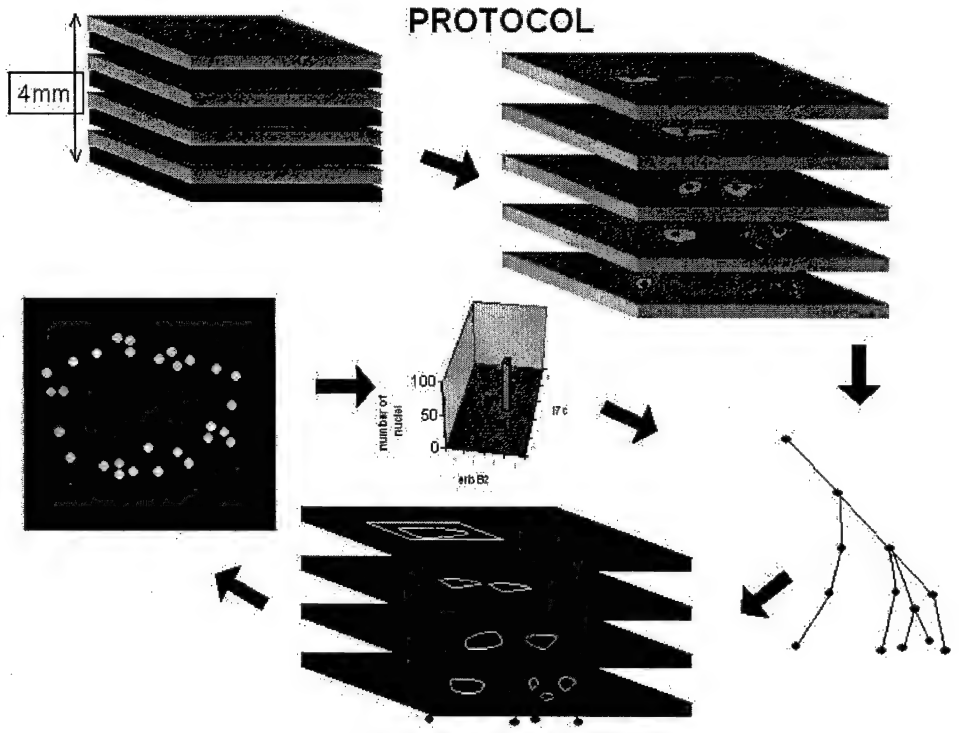
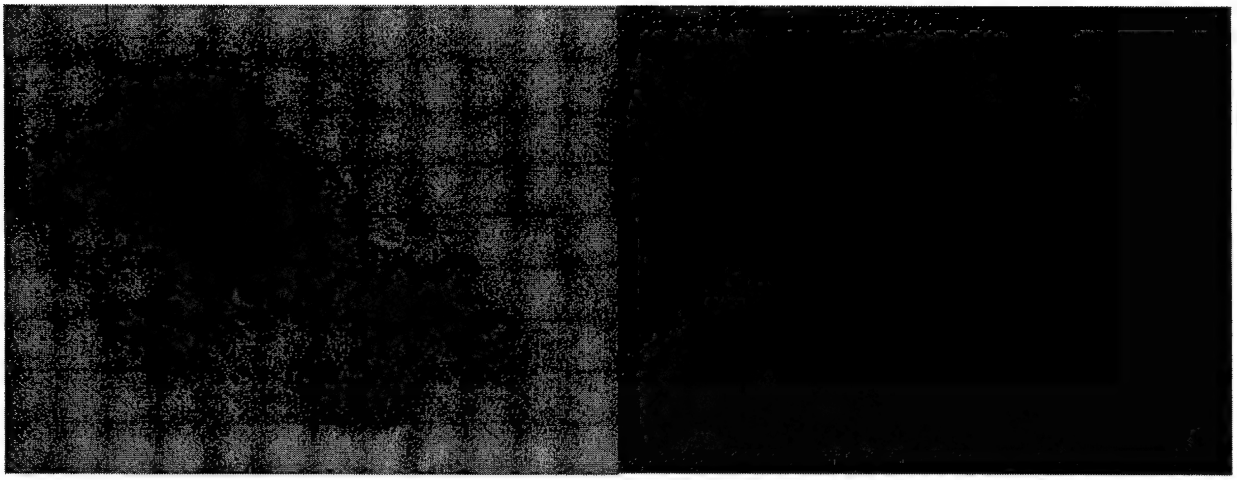



Figure 1: Protocol followed on tissue blocks. The different steps (sectioning, annotation, reconstruction, high magnification acquisition and molecular analysis) are illustrated.



a b

Figure 2: Background correction. a) Image of a section belonging to the human case. The background pattern created by the acquisition method is evident. b) Same image after background correction.

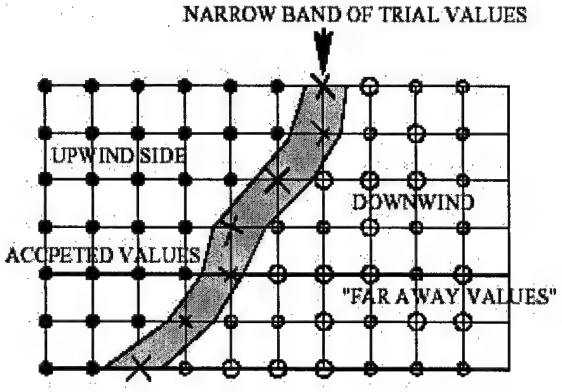



Figure 3: Snapshot of a 2D grid after several steps of the fast marching method. Accepted (*Visited*), narrow-band (*Close*) and far away (*Far*) elements are displayed.

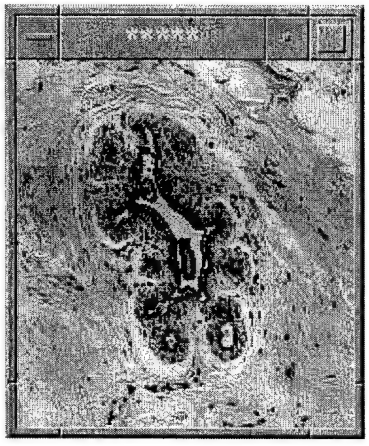


Figure 4: Result of the fast marching method on a duct in a human mammary gland tissue block. The dark point in the middle of the duct is the initial seed.

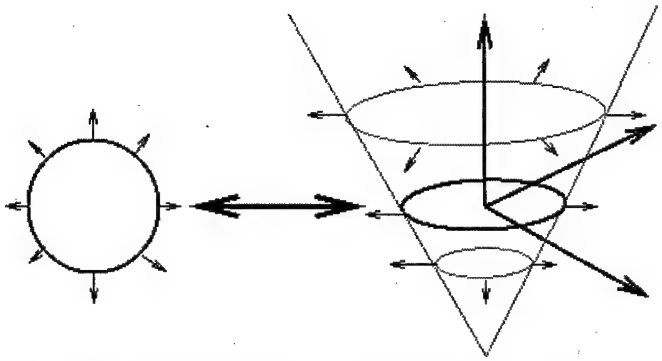


Figure 5: Basic idea of the level set method. The marching front is em bedded as the zero level set of a higher dimensional function.

Figure 6: Result of combining the fast marching and the level set methods on the same duct that we used in figure 4. The dark point in the middle of the duct is the initial seed.

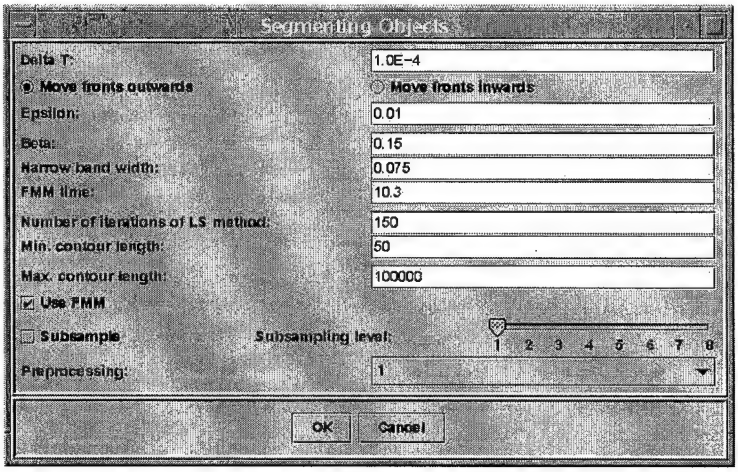
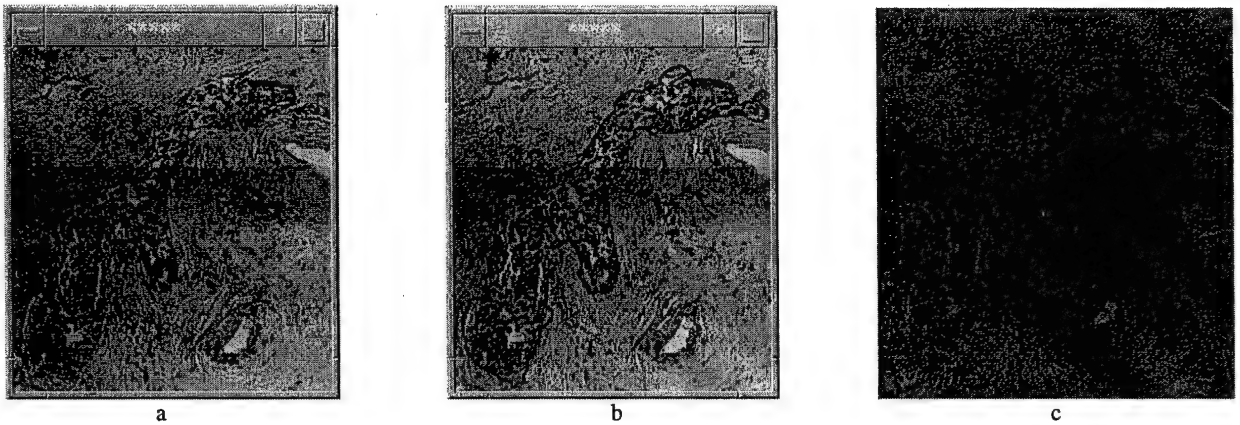
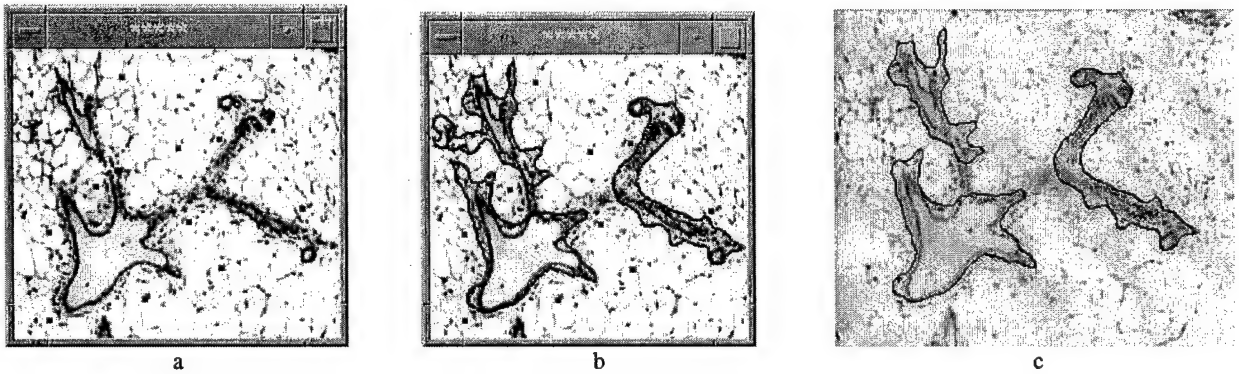



Figure 7: Options dialog. These are the parameters that the user can modify: *Delta T*, direction of front expansion, *Epsilon*, *Beta*, *Narrow band width* , *FMM time*, *Number of iterations of LS method*, *Min. contour length*, *Max. contour length*, *Use FMM*, *Subsample*, *Subsampling level* and *Preprocessing*.



a b c

Figure 8: Segmentation of a tum or in human mammary gland tissue. a) Initial "seed" points (blue dots). b) Results of the segmentation. Red lines delineate tumor masses. c) Results after adding the result to the original image and editing using the interactive tools provided by the syste m.



a b c

Figure 9: Segmentation of normal ducts in mouse mammary gland tissue. a) Initial "seed " points (blue dots). b) Results of the segmentation. Red lines delineate normal ducts. c) Results after adding the result to the original image and editing using the interactive tools provided by the system.

Figure 10: 3D reconstruction of tumors in a human mammary gland tissue block. Tumor masses are rendered as gray volumes. Notice the difference in smoothness between the first few sections in the central volume and the rest of the tumor due to automatic versus manual segmentation of the contours.

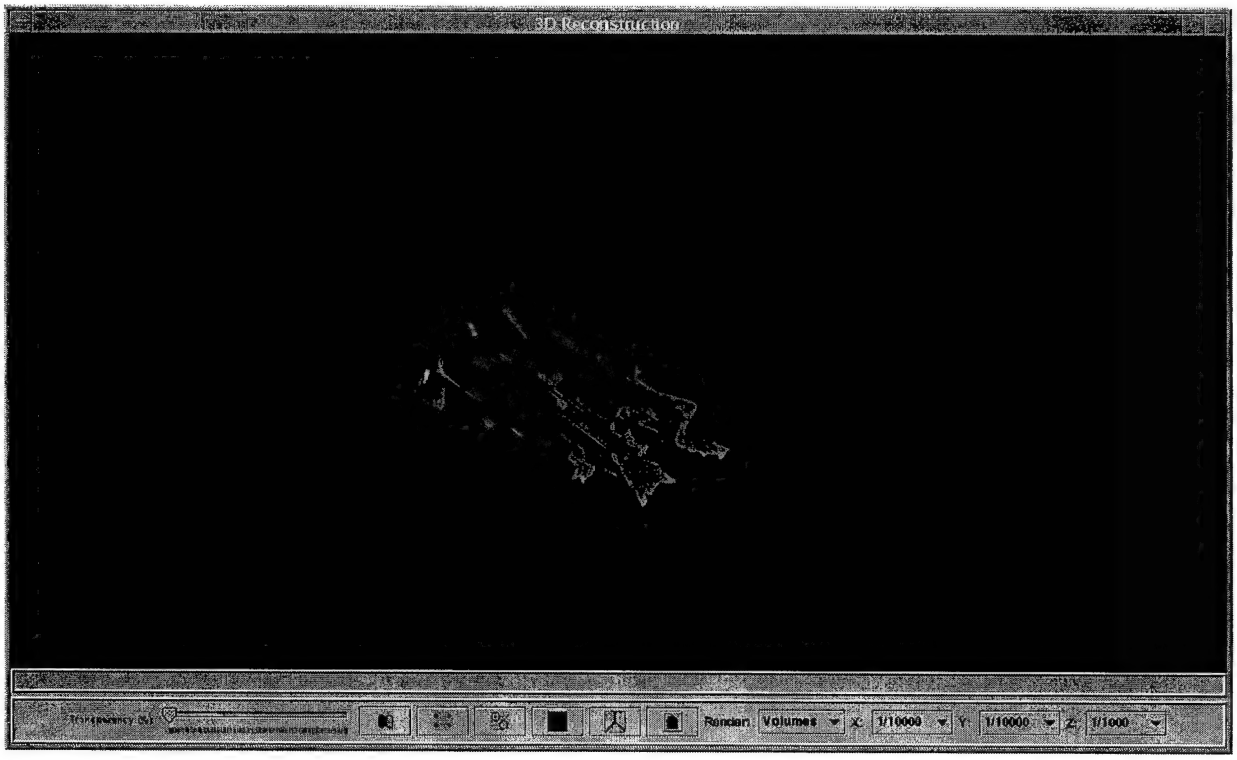



Figure 11: 3D reconstruction of normal ducts in a mouse mammary gland. Ducts are rendered as gray volumes. A single duct and its branches can be traced throughout the gland.

# Automatic segmentation of histological structures in mammary gland tissue sections

R. Fernandez-Gonzalez[1,4,*], T. Deschamps[2,3], A. Idica[1], R. Malladi[3] and C. Ortiz de Solorzano[1]
[1]Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley
[2]Mathematics Department, University of California, Berkeley
[3]Mathematics Department, Lawrence Berkeley National Laboratory, Berkeley
[4]UC Berkeley & San Francisco Joint Graduate Group in Bioengineering, Berkeley
{rfgonzalez,tdeschamps,akidica,r_malladi, codesolorzano}@lbl.gov

Real-time three-dimensional (3D) reconstruction of human mammary gland tissue blocks mapped with different markers of breast cancer is an extremely helpful tool for diagnosis, as well as for treatment planning and follow-up. This tool could also be utilized for the study of the molecular basis of breast cancer.

In this paper we present a framework for real-time segmentation of histological structures in images of normal and neoplastic mammary gland sections, acquired with the 3D microscopy system we have developed for reconstruction of thick tissue samples.

Paraffin embedded or frozen tissue blocks are first sliced, and sections are stained with Hematoxylin and Eosin. The sections are then imaged using conventional bright field microscopy and their background is corrected using a phantom image.

We then use the *Fast-Marching* algorithm to roughly extract the contours of the different morphological structures in the images. The result is then refined with the *Level-Set* method which converges to an accurate (sub-pixel) solution for the segmentation problem. Finally, our system stacks together the 2D results obtained in order to reconstruct a 3D representation of the entire tissue block under study. Our method is illustrated with results from human and mouse mammary gland tissue samples.

**Keywords:** Breast Cancer, Molecular Analysis, Automatic Segmentation, 3D Reconstruction, Fast-Marching, Level-Set

*Corresponding author: R. Fernandez-Gonzalez, Life Sciences Division, Lawrence Berkeley Lab, 1 Cyclotron Road, Building 84, MS 84-171, Berkeley, CA 94720, USA, Phone: (510) 486 5359, Fax: (510) 486 5730, e-mail: RFGonzalez@lbl.gov

The mature mammary gland has been characterized as a tree-like organ posessing three different branching levels and formed by hollow ducts converging at the nipple [1]. The ducts are lined by epithelial cells and terminate in secretory lobuloalveolar structures, which are the sites of milk production during lactation. In cancer, this hierarchical organization is disrupted by uncontrolled growth of the epithelium, resulting in the subsequent invasion of the surrounding tissue [2]. Thus, morphological or structural changes accompany the genetic changes that occur in breast cancer (see figure 1). An example of this correlation is seen in ductal carcinoma *in situ* (DCIS) which is a preinvasive lesion on which we focus many of our studies. In DCIS the molecular changes common in breast cancer can be observed together with a well-defined morphological pattern. Here, aneuploid nuclei or estrogen-receptor negative ones form hyperplastic lobules posessing an intact basement membrane and a necrotic center [3].

An interesting problem is the quantification of these genetic or molecular changes related to breast cancer progression in the context of the tissue environment where they occur. This quantification would allow the study of the correlation between morphological and genetic alterations in breast cancer. Since the mammary gland and its constituent elements are neither flat nor homogeneous, the approach to this problem should be three-dimensional as well, and take into account the heterogeneity of the gland. However, most classical methods in biology focus on only a single type of abnormalities (molecular or morphological) and neglect three-dimensionality and heterogeneity.

Although imaging of both tissue structure and function in vivo would be extremely desirable, the existing in vivo imaging methods (X-ray, MRI, optical tomography, etc.) do not provide the necessary resolution or penetration into the tissue for cell-level molecular analysis and/or require the use of an imaging-forming substrate that cannot be introduced in the tissue without severely altering its normal function. Consequently, ex-vivo microscopic analysis of the tissue from flat fixed sections is the routine method in histopathology. However, the limited capacity for people to extrapolate and visualize 3D structures from sequences of 2D scenes renders this method non suitable for quantitative 3D tissue characterization. To overcome these issues we have developed a system for simultaneous morphological and molecular analysis of thick tissue samples [4]. The system is composed of a computer-assisted microscope and a JAVA-based image display, analysis and visualization program (R3D2). R3D2 also allows semi-automatic acquisition, annotation, storage, 3D reconstruction and analysis of histological structures of interest (intraductal tumors, normal ducts, blood vessels, etc.) in thick tissue specimens. For this purpose the tissue needs to be embedded in a permanent or semi-rigid medium after collection. The tissue is then fully sectioned, and the resulting sections are stained in a way that visually highlights the desired structures. In histopathology, *Hematoxylin and Eosin* (H&E) is a common staining protocol to study the morphology of the tissue. In order to image not only structure, but also molecular events, we alternate H&E staining with fluorescence (immunofluorescence, fluorescence *in situ* hybridization,...) in consecutive sections.

This paper will focus on the annotation of the structures of interest on the H&E stained sections. Manual segmentation has been used before to delineate histological structures [5], [6]. In order to build the 3D reconstruction of the block, we initially annotated each of the interesting features of the tissue on each section by using manually drawn contours. This step constituted the bottleneck in the study of samples. Semi-automatic approaches to segmentation of features of interest in histological sections have also been used [7], but they still involve too much user interaction, which remains the most time intensive step in the sample analysis process. Automatic annotation (segmentation) of the structures of interest is the answer to this problem. In this paper, we propose an automatic method followed by interactive correction that greatly reduces the amount of interaction required, thus eliminating the bottleneck and allowing us to use our system for imaging and reconstruction of large, complex tissue structures.

Automatic extraction of contours in 2D images is usually done with *Active contour models*, originally presented by *Kass, Witkins and Terzopoulos* [8]. These methods are based on deforming an initial contour (polygon) toward the boundary of the desired object to extract in an image. The deformation is achieved by minimizing a certain energy function, which is computed by integrating along the contour terms related to its continuity/derivability and terms related to the pixel values of the area of the image where the contour is defined. That energy function approaches a minimum near the object boundary, and thus, the minimization process drives the curve toward the desired shape.

As an alternative, implicit surface evolution models have been introduced by Malladi *et. al* [?], [9] and Caselles *et. al* [?]. In these models, the curve and surface models evolve under an implicit speed law containing two terms, one that attracts it to the object boundary and the other that is closely related to the regularity of the shape. Specifically, the proposal is to use the *Level-Set* approach of *Osher and Sethian* [10]. This is an interface propagation technique used for a variety of applications, including segmentation. The initial curve is represented here as the zero *Level-Set* of a higher dimensional function, and the motion of the curve is embedded within the motion of that higher dimensional function. An energy formulation similar to the *Active Contours* leads to a minimization process with several advantages. First, the zero *Level-Set* of the higher dimensional function is allowed to change topology and form sharp corners. Second, geometric quantities such as normal and curvature are easy to extract from the hypersurface. Finally, the method expands straightforwardly to 3D [11], but adding a dimension to the problem increases the computational cost associated with the method. The narrow-band approach of *Adalsteinsson and Sethian* [12] accelerates the *Level-Set* flow by considering for computations only a narrow band of pixels around the zero level-set. But in our experience, the narrow band technique does not reduce the computational cost to a reasonable limit, due to the large size of the images of the sections.

Thus, we propose to use the *Fast-Marching* method [13]. This method considers monotonically advancing fronts (speed always positive or negative), providing a result very quickly, albeit one that is not as accurate as the one obtained by

using the *Level-Set* algorithm. *Malladi and Sethian* [14] showed that the *Fast-Marching* method can be used as the initial condition for the slower but more accurate *Level-Set* segmentation, obtaining real-time delineation of the structures of interest. The combination of all these tools is the framework we are using in order to reconstruct normal and cancerous ducts in mammary gland tissue sections.

This paper is organized as follows: section I describes the general tissue handling and image acquisition protocols that we use, as well as the theoretical basis of the segmentation scheme; section II shows the results of applying our method to histological tissue sections; and section III discusses the results and suggests some future developments and improvements to our approach.

## I. Methodology

### *A. Tissue processing and imaging*

The tissue processing and staining protocol used is illustrated in figure 2.

Tissue blocks of 4 to 5 mm thickness are sliced into $5\mu m$ (thin) sections (Step 1). The odd sections are stained with H&E,a technique routinely used in histopathology to obtain morphological information both at the cytological (single cell) and architectural (organ) levels. The even sections are stained using some fluorescence technique (immuno-cytochemistry, fluorescence in situ hybridization, ...) depending on the molecular phenomena that we want to study. Describing the acquisition and analysis of the fluorescent images is out of the scope of this paper. Therefore, the rest of this section describes only the protocol used for the H&E stained sections. Low magnification ($2.5\times$) images of all the sections are automatically acquired using a motorized *Zeiss Axioplan I* microscope coupled with a monochrome *XilliX Microimager* CCD camera (Step 2): the area of the slide occupied by the tissue is scanned in a process that takesone image per field of view. Then all the individual snapshots are tiled together into a single, whole-view image of the sections. The sequence of microscope movements and camera operations is determined by an application running on the *Sun Ultra 10* workstation that controls the camera and all moving parts of the microscope.

The next step consists of annotating the structures of interest (ducts, lymph nodes, tumors, ...) in the images of the H&E stained sections (Step 3). These annotated structures are then used to render a three-dimensional model of the tissue block (Step 4) where the morphology of the sample helps in determining which are the areas of the tissue where a molecular analysis may be more interesting. After selecting these areas in the three-dimensional rendition of the organ, the system asks the user to place the corresponding fluorescent section(s) on the stage (Step 5), and high magnification ($40 - 100\times$) images of the chosen areas are acquired (Step 6). Finally, quantification routines are run on these new images and results are collected (Step 7).

Manually annotating morphological structures in complex sections of tissue blocks containing dozens of sections at the

very least is feasible, but for all purposes impractical due to the tremendous human effort involved. Although it might be the most accurate and reliable approach, manual annotation is not possible to carry out except when reconstructing small, simple tissue volumes. As a result, we have developed automatic methods that eliminate or greatly reduce human interaction, thus making the reconstruction of complex systems possible. What follows describes our approach.

*B. Segmentation*

B.1 Background removal

The algorithm that we use to acquire an image of an entire section creates a mosaic from a set of snapshots (one per field of view). This approach gives rise to a background pattern across the image (figure 3-left) involving relatively large gradients in between elements of the mosaic.

Objects of interest often span through several fields of view, and since our segmentation approach depends largely on the gradients of the image, we need to eliminate the background pattern in order to obtain good segmentation results.

This can be done by performing a set of arithmetic operations on the "mosaic" image, known as *background compensation.* First we need a *phantom*, that is, an image of an empty field of view taken under the same illumination conditions and microscope configuration that we used to acquire the initial image. Since most of the images that we acquire have an empty frame in the upper left corner, it is simple to choose that frame as our *phantom* for the corresponding section. After normalizing the pixel values in the *phantom*, and for each frame in the entire image, we divide the value at each pixel by the value at the corresponding pixel in the *phantom* frame. The resulting image is background-corrected as shown in figure 3-right, and it is a better input for our segmentation algorithms.

B.2 Preliminary segmentation: Fast-Marching method

Consider a monotonically advancing 2D front $C$ with speed $F$ always positive in the normal direction, starting from an initial point $\mathbf{p}_0$,

$$\frac{\partial C}{\partial t} = F\vec{n}. \tag{1}$$

This equation drives the evolution of a front starting from an infinitesimal circular shape around $\mathbf{p}_0$ until each point $\mathbf{p}$ inside the image domain is visited an assigned a crossing time $U(\mathbf{p})$, which is the time $t$ at which the front reaches the point $\mathbf{p}$.

The gradient of the arrival time is inversely proportional to the speed function, and thus we have a form of the Eikonal equation

$$|\nabla U|F = 1 \text{ and } \quad U(\mathbf{p}_0) = 0 \tag{2}$$

Classic finite difference schemes for Equation (2) tend to overshoot and are unstable. An up-wind scheme was proposed

by *Sethian* [11]. It relies on a one-sided derivative that looks in the up-wind direction of the moving front, and thereby avoids the over-shooting associated with finite differences:

$$\left(\max\left(u - U_{i-1,j}, u - U_{i+1,j}, 0\right)\right)^2 + \left(\max\left(u - U_{i,j-1}, u - U_{i,j+1}, 0\right)\right)^2 = \frac{1}{F_{i,j}^2} \tag{3}$$

giving the correct viscosity-solution $u$ for $U_{i,j}$.

One approach to obtain these solutions is to use the *Fast Marching* method. It introduces an order in the way in which the points in the grid are selected, based on the fact that information is propagating *outward*, since the front can only grow due to Equation (3) -which is quadratic-. Dijkstra introduced the approach of considering only the necessary set of grid points at each step. This method, originally used for the construction of minimum length paths in a graph between two given nodes [15], is also used in our algorithm in order to reduce computational cost. The algorithm is detailed in Table I.

Notice that in solving Equation (3), only *Alive* points are considered. This means that for each point the calculation is made using the current values of $\mathcal{U}$ at the neighbors and not estimates, $U$. at other trial points. Considering the neighbors of the grid point $(i, j)$ in 4-connectedness, we designate $\{A_1, A_2\}$ and $\{B_1, B_2\}$ as the two couples of opposite neighbors such that we get the ordering $U(A_1) \leq U(A_2)$, $U(B_1) \leq U(B_2)$, and $U(A_1) \leq U(B_1)$. Since we have $u \geq U(B_1) \geq U(A_1)$, we can derive

$$(u - U(A_1))^2 + (u - U(B_1))^2 = \frac{1}{F_{i,j}^2} \tag{4}$$

Computing the discriminant $\Delta$ of Equation (4) we complete the steps described in table II. Thus the algorithm needs only one pass over the image to find a solution. To execute all the operations that we just described in the minimum amount time, the *Trial* points are stored in a min-heap data structure [11]. Since the complexity of changing the value of one element of such a heap is bounded by a worst-case bottom-to-top processing of the tree ($O(\log_2 N)$), the total algorithm has a complexity of $O(N \log_2 N)$ on a grid with $N$ nodes.

Finally, we define the speed of propagation in the normal direction as a decreasing function of the gradient $|\nabla I(\mathbf{x})|$, that is, a function that is very small near large image gradients (i.e. possible edges) and large when the brightness level is constant:

$$F(\mathbf{x}) = \exp -\alpha |\nabla I(\mathbf{x})| \ , \ \alpha > 0 \tag{5}$$

where $\alpha$ is the *edge strength*, or the weight that we give to the presence of a gradient in order to slow down the front. Depending on this value the speed function falls to zero more or less rapidly, and thus, it could stop a few grid points away from the real edge. Also, variations in the gradient along the boundary can cause inaccurate results. False gradients due to noise can be avoided using an edge preserving smoothing scheme on the image as a preprocessing step.

The user can run the *Fast-Marching* flow from a given set of initial points (mouse clicks on the background of the images). Alternatively, the user can decide to segment only the structures within a manually defined rectangular region of interest. If the region is too big, sub-sampling can be used so that the segmentation process is not too slow. However, this option must be used carefully, since sub-sampling smooths the boundaries of the objects present in the image and can completely obliterate smaller structures. The resulting contour (or contours if we are segmenting several objects at the same time) will provide an excellent initial condition for the *Level-Set* method.

Putting all of these elements together we are able to get a good approximation of the shape of the object that we are trying to segment (figure 5).

In order to improve the final result, we propose to run a few iterations of the *Level-Set* method using the result of the *Fast-Marching* method as the initial condition.

### B.3 Final segmentation: *Level-Set* method

Once we have obtained a good approximation of the shape of the object using the *Fast-Marching* algorithm, we can afford to use the more computationally expensive *Level-Set* method to improve the result of the segmentation. The essential idea here is to embed our marching front as the zero *Level-Set* of a higher dimensional function. In our case, we take that function to be $\phi(\mathbf{x}) = \pm d$, where $d$ is the signed distance from $\mathbf{x}$ to the front (see figure 4), assigning negative distances for pixels inside the evolving curve, and positive distances for pixels outside it.

Thus, at time $T$, the front sits exactly at the intersection of $\phi$ with the $XY$ plane (zero *Level-Set* of $\phi$). The rest of the Level-Sets are isodistance curves. Using the chain rule we can find the motion equation:

$$\phi_t + F(x,y)|\nabla\phi| = 0 \tag{6}$$

where $F$ is again the speed in the normal direction. This is an initial value partial differential equation, since it describes the evolution of the solution on the basis of an initial condition defined as $\phi(\mathbf{x}, t = 0) = \phi_0$. As pointed out before, the *Level-Set* approach offers several advantages:

- the zero *Level-Set* of the function can change topology and form sharp corners;
- a discrete grid can be used together with finite differences to approximate the solution;
- intrinsic geometric quantities like normal and curvature can be easily extracted from the higher dimensional function, and
- everything extends directly to 3D.

To mold the initial condition (in our case the result of the *Fast-Marching* method) into the desired shape we use two force terms. By substituting these two terms in the motion equation [16] we get:

$$\phi_t - g(1 - \epsilon\kappa)|\nabla\phi| - \beta\nabla g \cdot \nabla\phi = 0 \ , \epsilon > 0 \ , \beta > 0 \tag{7}$$

where $g$ is an edge indicator function defined by the speed of the front (5) in the *Eikonal* Equation (2); $\kappa$ is the curvature of the expanding front, and $\phi_t$ is the unknown that we are trying to compute. As before, the image $I(\mathbf{x})$ can be preprocessed using an edge-preserving smoothing scheme.

The second term of the equation has two components. The first one attracts the surface toward high gradients (edges), while the second one (*motion by curvature*) opposes this inflationary movement: it accelerates those parts of the contour where the curvature is negative (they are behind the average position of the front) and slows down the parts of the contour with positive curvature (advanced to the rest of the curve). The parameter $\epsilon$ is the weight of the *motion by curvature* term, and determines the strength of its regulatory effect: the bigger we make $\epsilon$, the more we constrain the possibility to obtain sharp corners and irregular contours. In practice, an intermediate value of $\epsilon$ provides a good trade-off between contour smoothing and accuracy.

Finally, the last term of the equation adjusts the front to the object boundaries once it has found them by the effect of the previous term. It aligns all the Level-Sets with the ideal gradient, which would be a step function centered at the point of maximum gradient in the original image. $\beta$ is the weight of the advection of the front by the edge vector field $\nabla g$. It determines the strength of the attraction of the front to the edges.

At times, for very small objects, it is possible to use the *Level-Set* method from the initial point. However, for any type of morphological structure, we can use the result of the *Eikonal* equation (2), $U(x, y)$, as the initial condition for the *Level-Set* algorithm: $\phi(x, y; t = 0) = U(x, y)$. Then by solving Equation (7) for a few time steps using the narrow band approach, we obtain an accurate, real-time segmentation of the desired object (see figure 5).

## II. Results

In this section we consider the problem of reconstructing DCIS areas in a tissue biopsy of a cancerous human breast as well as a group of ducts through an entire mouse mammary gland. The tissue samples were sliced, and we used 55 sections for the human case and 40 for the mouse one. Manually delineating each one of the structures in every section is an extremely time-consuming process. To automatically segment those structures using the framework described in section I, we begin by defining a Region-Of-Interest (*ROI*) where we will run the segmentation algorithm. This *ROI* can be extended to cover the entire section, but considering the size of the images it is wise to use sub-sampling in order to run the algorithm in real time. The level of sub-sampling can be determined by the user: greater sub-sampling can be used on large *ROIs*, without compromising the resolution and accuracy of the final segmentation.

After defining the *ROI*, we have to tune the different parameters of the segmentation process, particularly $\alpha$ (Equa-

tion (5)) and $\epsilon$ (Equation (7)). This is done by the user based on the default values provided by the algorithm and the type of object that he/she is trying to segment. However, we have observed that a particular set of parameters is frequently good enough to segment similar structures (i.e., all the tumors, all the ducts, all the lymph nodes, ...) throughout all the sections of a particular tissue block. Thus, the user only needs to modify the parameters the first time that he/she tries to segment a new type of morphological element in the tissue.

After selecting the parameters of the flow, initial points are defined inside (to find the internal contour) or outside (to find the external contour) of the structures of interest. In most cases one mouse-click is enough, though large images may require several, evenly distributed clicks. The value of $U(\mathbf{x})$ at these points is set to zero as in equation 2, and the *Fast-Marching* algorithm of table I is executed. When this method finishes, the final $U(\mathbf{x})$ function is passed as the initial condition to the *Level-Set* motion equation (7). We then iterate this equation for a few steps. This segmentation scheme provides a result in less than one second for images whose size (after sub-sampling, if any) is around 2 kilobytes (e.g., $512 \times 512$ pixels) running on a *Sun Ultra 10* workstation with 1 GB of RAM.

### *A. Human case segmentation*

Figure 6 displays an example of the result obtained with the combination of the *Fast-Marching* and the *Level-Set* methods in a tissue biopsy of a cancerous human breast.

Figure 6-middle shows the segmentation of a tumor mass in a human tissue block. The results of the segmentation can be edited and removed with the interactive tools provided by our system (see figure 6-right). For this segmentation, an area was selected around the structures of interest and no sub-sampling was used. The initial contours are represented by blue points in figure 6-left.

### *B. Mouse case segmentation*

In figure 7 we can see the segmentation of the external contours of several ducts in a particular area of a mouse mammary gland.

In this case we also run the algorithm on a *ROI* on one of the sections with no sub-sampling factor. Figure 7-left displays the data where we initialized the initial contours. Figures 7-middle and right show the results of the segmentation before and after interactive correction of the results respectively.

### *C. 3D Reconstruction*

Finally the segmented shapes are connected (manually) between sections, and 3D reconstructions of the samples are built. Figure 8 shows a reconstruction of the tumors contained in the human tissue block, including the tumor shown in figure 6.

Increasing the "motion by curvature" term, as described in section I, can reduce surface noise. In figure 9 we can see the reconstruction of the normal ducts segmented in the images of the mouse mammary gland (see one of the corresponding 2D segmentations in figure 7).

*D. Further examples*

Figures 10 and 11 show two more examples of the results that can be obtained with our segmentation approach. In both cases the initial *ROI* was sub-sampled by a factor of two in both the X and Y directions. Interestingly, the same segmentation parameters ($\alpha$ and $\epsilon$) were used for both examples.

Figure 10-left shows an DCIS lesion together with a normal duct in a human tissue section. Both structures were segmented at the same time using a single initial point, as can be seen in figure 10-middle. Figure 10-right shows the results incorporated to the full resolution image.

In figure 11-left, a terminal ductal lobular unit (TDLU) can be observed. These are lobuloalveolar structures where milk is produced during lactation in the human breast. The multiple alveoli that form the TDLU, together with the presence of a ductal part, make automatic segmentation of this type of structure a difficult task. However, after subsampling the image, our algorithm is able to find a contour that surrounds the entire structure (figures 11-middle and 11-right), thus rendering possible its reconstruction.

## III. DISCUSSION

We have developed a microscopy system that combines 3D reconstruction of normal and diseased mammary gland specimens with their morphological and molecular analysis. However, the interaction required to operate the system is quite intensive, limiting the scope of its application to studies not requiring a high throughput in the analysis of samples. In this paper we have presented a method that reduces the time and interaction needed to build the 3D model of a tissue block, enormously increasing the potential throughput of the system and therefore allowing us to use this approach for the analysis of large, complex specimens. To achieve this goal we have combined image processing techniques and two well-established schemes for interface propagation: the *Fast-Marching* method and the *Level-Set* method.

Our approach starts by correcting the background of the images. This is an important step, since the background pattern generated during image acquisition modifies the gradient of the image, and the speed function that we use for interface propagation depend on that gradient. Once the background has been corrected, we run the *Fast-Marching* method. This technique provides a good approximation of the boundaries of the objects that we are trying to segment in a very short time, since it assumes monotonic speed functions (always positive or always negative). We then use the approximation provided by the *Fast-Marching* method as the initial condition for the *Level-Set* method. This more

computationally expensive algorithm is run for just a few steps, enough to fit the front to the contours of the structures of interest, but not as many as to make the segmentation too time consuming.

Though this approach is very useful, it can still be improved. The most accurate segmentations (and reconstructions) are obtained on full resolution images. However, for some large structures like lymph nodes, or when trying to delineate multiple elements at the same time (for example a group of ducts), segmentation on a full resolution image is not real-time any more, and can take up to one minute. Using sub-sampling takes the segmentation execution time back to real time at the expense of some accuracy loss. Also, in areas with a lot of stromal texture tuning the parameters of the algorithm ($\alpha$ and $\epsilon$) becomes more difficult. At times this process can take a few trials, since the expanding front tends to get stuck in high gradients that do not correspond to the boundaries of the feature that we are trying to segment, but to the texture of the stroma. Finally, once all the structures of interest have been segmented, you still need to manually connect them between sections. This constitutes a new bottleneck in the tissue analysis process.

For these reasons, we are currently working on a time-step independent scheme that is expected to be faster than the current one. To improve the accuracy of the results, we have developed an edge-preserving smoothing algorithm based on the *Beltrami* flow [17], which can replace the Gaussian smoothing currently used before executing the segmentation methods. This algorithm eliminates false gradients due to noise, while enhancing gradients due to object boundaries, thus allowing the front to fit the boundaries of the object more accurately. Finally, the *Level-Set* approach is readily extensible to 3D. The ability to segment 3D structures of interest versus 2D ones would save the process of connecting the segmented 2D contours from section to section, thus improving the analysis time. Also, since geometric properties can be easily extracted from the higher dimensional function used in the *Level-Set* algorithm, we could readily obtain some information about the extracted volume from the segmentation algorithm itself.

In conclusion, we have presented a real-time method for automatic segmentation of morphological structures in mammary gland tissue sections. It is precisely the delineation of those structures that required heaviest user interaction in our sample-analysis protocol. Therefore, the automatic approach to segmentation that we describe here represents a first step toward real-time reconstruction and analysis of mammary gland samples. Achieving that goal would allow us to accelerate our studies on the biological basis of human breast cancer. Moreover, obtaining real-time reconstruction and analysis of samples from our system would be useful for pathological diagnosis in a clinical environment: 3D renderings of all the morphological structures in a mammary gland biopsy could be mapped with the distribution of particular markers of breast cancer within few hours of extracting the tissue from the patient. From an intra-surgically point of view, the renderings would prove - tissue processing and stain permitting - as an important tool in the evaluation of breast tumors and their margins.

## Acknowledgments

The U.S. Army Medical Research Materiel Command under grants DAMD17-00-1-0306 and DAMD17-00-1-0227 and the Lawrence Berkeley National Laboratory Directed Research and Development program supported this work. This work was also supported by the Director, Office of Science, Office of Advanced Scientific Research, Mathematical, Information, and Computational Sciences Division, U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

## References

[1] L. Hennighausen and G.W. Robinson, "Think globally, act locally: the making of a mouse mammary gland," *Genes Development*, vol. 12, no. 4, pp. 449–455, february 1998.

[2] L.M. Franks and N.M. Teich, *Introduction to the cellular and molecular biology of cancer*, Oxford University Press, Oxford, 1997.

[3] T. Tot, L. Tabar, and P.B. Dean, *Practical Breast Pathology*, Thieme, Falun Central Hospital, 1st edition, 2002.

[4] R. Fernandez-Gonzalez, A. Jones, E. Garcia-Rodriguez, P.Y. Chen, A. Idica, M.H. Barcellos-Hoff, and C. Ortiz de Solorzano, "A system for combined three-dimensional morphological and molecular analysis of thick tissue samples," *Microscopy Research and Technique*, vol. 59, no. 6, pp. 522–530, december 2002.

[5] D.F. Moffat and J.J. Going, "Three-dimensional anatomy of complete duct systems in human breast: pathological and developmental implications," *Journal of Clinical Pathology*, vol. 49, pp. 48–52, 1996.

[6] T. Ohtake, I. Kimijima, T. Fukushima, M. Yasuda, K. Sekikawa, S. Takenoshita, and R. Abe, "Computer-assisted complete three-dimensional reconstruction of the mammary gland ductal/lobular systems," *Cancer*, vol. 91, no. 12, june 2001.

[7] F. Manconi, R. Markham, G. Cox, E. Kable, and I.S. Fraser, "Computer-generated, three-dimensional reconstruction of histological parallel serial sections displaying microvascular and glandular structures in human endometrium," *Micron*, vol. 32, pp. 449–453, 2001.

[8] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.

[9] R. Malladi, J.A. Sethian, and B.C. Vemuri, "Shape modelling with front propagation: A level set approach," *IEEE Transactions On Pattern Analysis And Machine Intelligence*, vol. 17, no. 2, pp. 158–175, Feb. 1995.

[10] S. Osher and J.A. Sethian, "Fronts propagating with curvature dependent speed: algorithms based on the hamilton-jacobi formulation," *Journal of Computational Physics*, vol. 79, pp. 12–49, 1988.

[11] J.A. Sethian, *Level set methods: Evolving Interfaces in Geometry, Fluid Mechanics, Computer Vision and Materials Sciences*, Cambridge University Press, University of California, Berkeley, 2nd edition, 1999.

[12] D. Adalsteinsson and J.A. Sethian, "A fast level set method for propagating interfaces," *Journal of Computational Physics*, vol. 118, pp. 269–277, 1995.

[13] J.A. Sethian, "A fast marching level set method for monotonically advancing fronts," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, no. 4, pp. 1591–1595, Feb. 1996.

[14] R. Malladi and J.A. Sethian, "A real-time algorithm for medical shape recovery," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV'98)*, Jan. 1998, pp. 304–310.

[15] E.W. Dijkstra, "A note on two problems in connection with graphs," *Numerische Mathematic*, vol. 1, pp. 269–271, 1959.

[16] C. Ortiz de Solorzano, R. Malladi, S.A. Lelievre, and S.J. Lockett, "Segmentation of nuclei and cells using membrane related protein markers," *journal of Microscopy*, vol. 201, no. 3, pp. 404, 2001.

[17] N. Sochen, R. Kimmel, and R. Malladi, "A general framework for low level vision," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 310–318, march 1998.

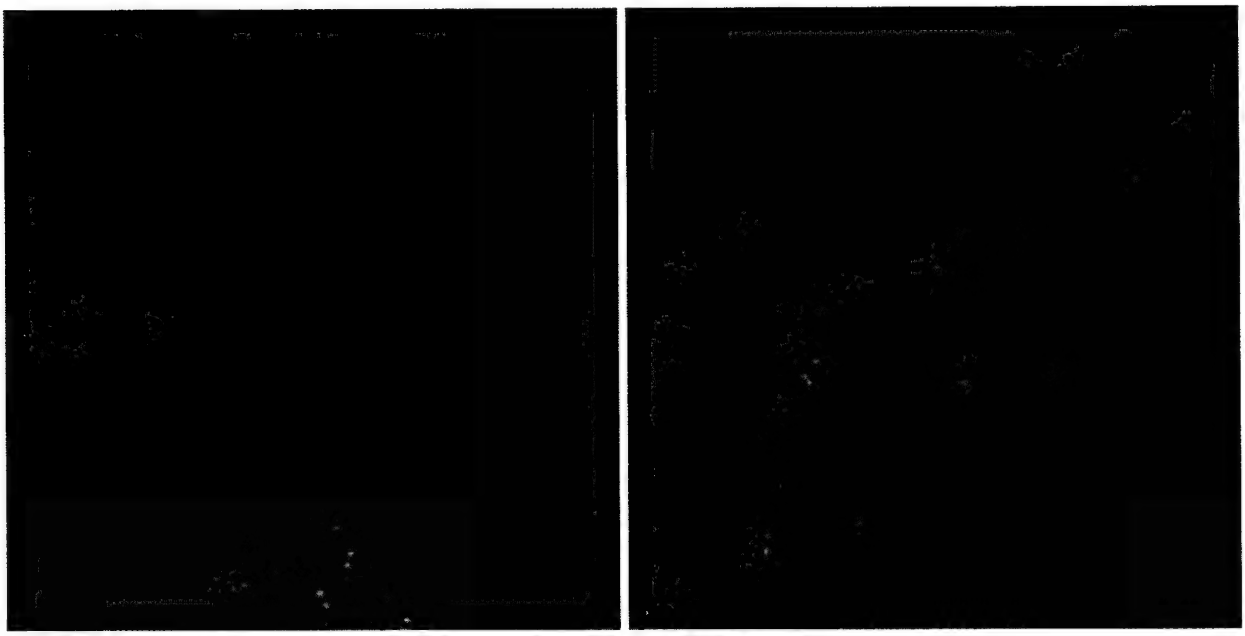


Fig. 1. Morphological and genetic alterations in breast cancer. The images show two optical sections of human mammary gland tissue acquired using a confocal laser scanning microscope; nuclei are displayed in red, a probe for a certain DNA sequence in chromosome 17 is shown in green. Left image: normal tissue; as expected, each nucleus contains up to two green signals (up to two copies of chromosome 17 per nucleus in a single optical section). Right image: neoplastic lesion; not only do some nuclei contain more than 2 copies of the probe, but also they have distinct morphological changes that can be observed in this section.

**Algorithm for 2D Fast Marching**

- Definitions:
  - *Alive* set: all grid points where the action value $U$ has been reached and will not be changed;
  - *Trial* set: next grid points (4-connectedness neighbors) to be examined. An estimate $U$ of $U$ has been computed using Equation (3) from *Alive* points only (i.e. from $U$);
  - *Far* set: all other grid points, where there is no estimate for $U$ yet;
- Initialization:
  - *Alive* set: reduced to the starting point $\mathbf{p_0}$, with $U(\mathbf{p_0}) = U(\mathbf{p_0}) = 0$;
  - *Trial* set: reduced to the four neighbors $\mathbf{p}$ of $\mathbf{p_0}$ with initial value $U(\mathbf{p}) = \frac{1}{F}(\mathbf{p})$ $(U(\mathbf{p}) = \infty)$;
  - *Far* set: all other grid points, with $U = \infty$;
- Loop:
  - Let $\mathbf{p} = (i_{min}, j_{min})$ be the *Trial* point with the smallest action $U$;
  - Move it from the *Trial* to the *Alive* set (i.e. $U(\mathbf{p}) = U_{i_{min},j_{min}}$ is frozen);
  - For each neighbor $(i, j)$ (4-connectedness in 2D) of $(i_{min}, j_{min})$:
  - If $(i, j)$ is *Far*, add it to the *Trial* set and compute $U_{i,j}$ using Eqn. 3;
  - If $(i, j)$ is *Trial* , update the action $U_{i,j}$ using Eqn. 3.

TABLE I

FAST MARCHING ALGORITHM

1. • If $\Delta \geq 0$, $u$ should be the largest solution of Equation (4);
   - If the hypothesis $u > U(B_1)$ is wrong, go to 2;
   - If this value is larger than $U(B_1)$, this is the solution;
   - If $\Delta < 0$, $B_1$ has an action too large to influence the solution. It means that $u > U(B_1)$ is false. Go to 2;

Simple calculus can replace case 1 by the test:

If $\frac{1}{F_{i,j}} > U(B_1) - U(A_1)$, $u = \frac{U(B_1)+U(A_1)+\sqrt{2\frac{1}{F_{i,j}^2}-(U(B_1)-U(A_1))^2}}{2}$ is the largest solution of Equation (4) else go to 2;

2. Considering that we have $u < U(B_1)$ and $u \geq U(A_1)$, we finally have $u = U(A_1) + \frac{1}{F_{i,j}}$.

TABLE II

SOLVING LOCALLY THE UP-WIND SCHEME

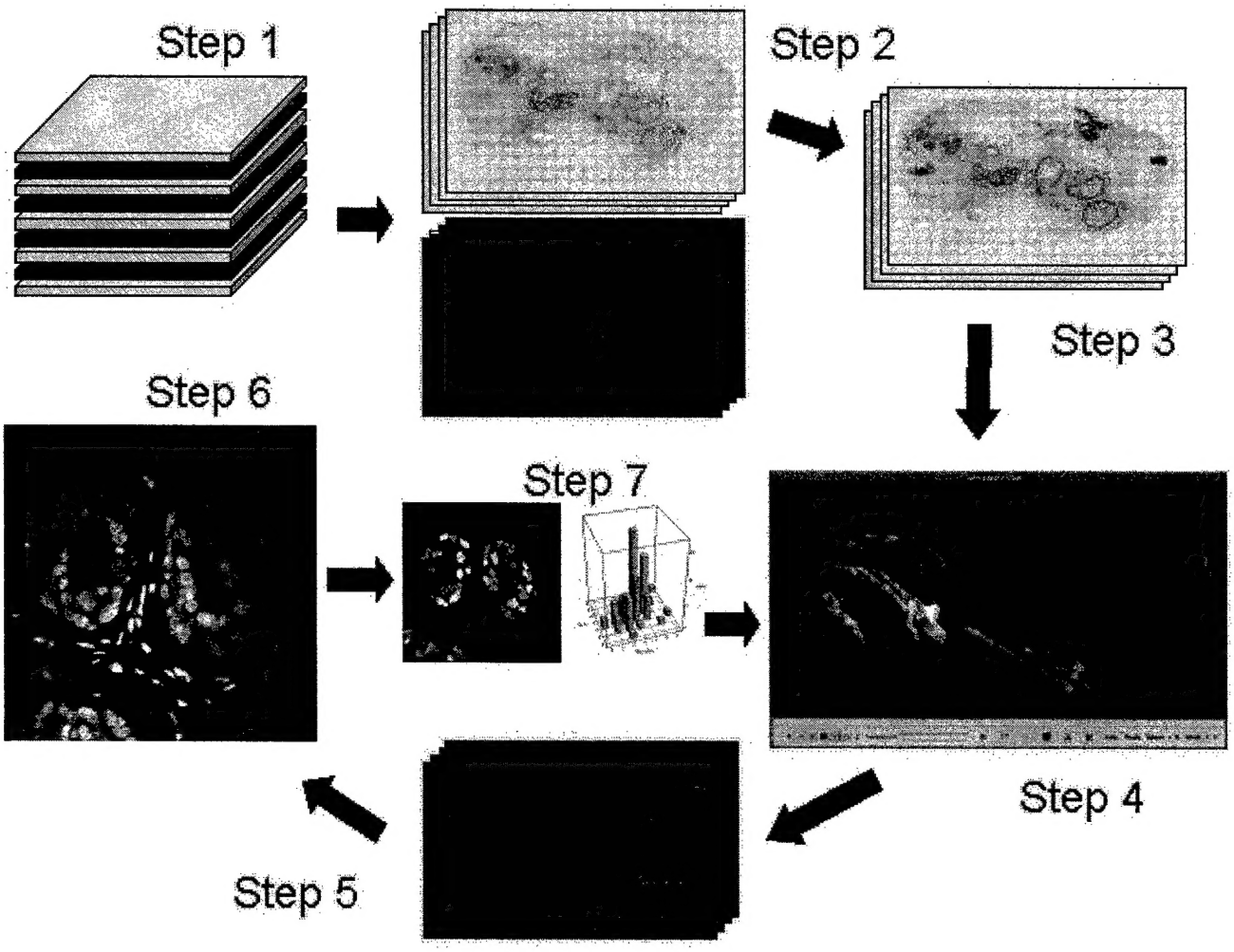



Fig. 2. Protocol followed on tissue blocks. The different steps (sectioning, annotation, reconstruction, high magnification acquisition and molecular analysis) are illustrated. Samples are fully sectioned at $5\mu m$ (Step 1). The odd sections are stained with H&E, the even ones with some kind of fluorescence technique (application-dependent). Images are acquired of all the sections (Step 2), and structures of interest are delineated in the H&E-stained ones (Step3). A 3D reconstruction of the specimen is created from these markings (Step 4). From the 3D reconstruction of the tissue different areas can be selected for molecular analysis (Step 5). The system will take high magnification images of those areas on the corresponding fluorescent sections (Step 6). Image analysis tools can then be used to quantify the presence and distribution of molecular markers in the high magnification images (Step 7).

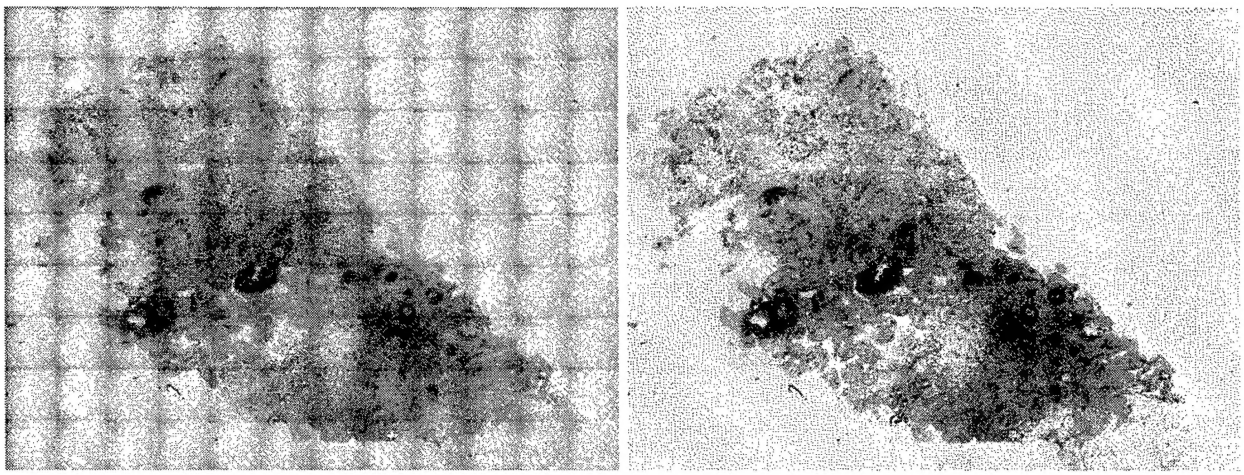


Fig. 3. Background correction. a) Image of a section belonging to a human case. The background pattern created by the acquisition method is readily noticeable. b) Same image after background correction.

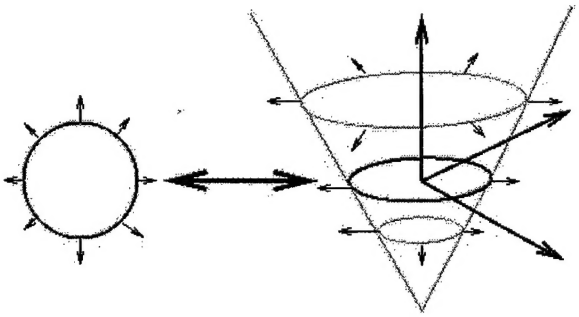


Fig. 4. Basic concept behind the *Level-Set* method. The marching front is embedded as the zero *Level-Set* of a higher dimensional function.

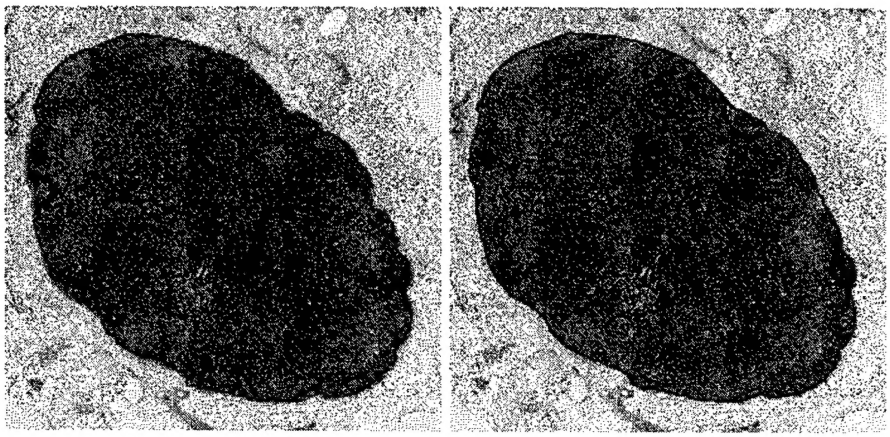


Fig. 5. The segmentation of a lymph node in a mouse mammary gland section is shown in black. Left image: result of the *Fast-Marching* method; it provides a good approximation to the boundaries of the lymph node; the blue point in the middle of the lymph node is the initial contour from which the algorithm was run. Right image: result of using the *Level-Set* method after the *Fast-Marching* method; the final contour is more accurate and smoother. In both cases the images were subsampled in the x and y directions to be able to run the segmentation in real time.

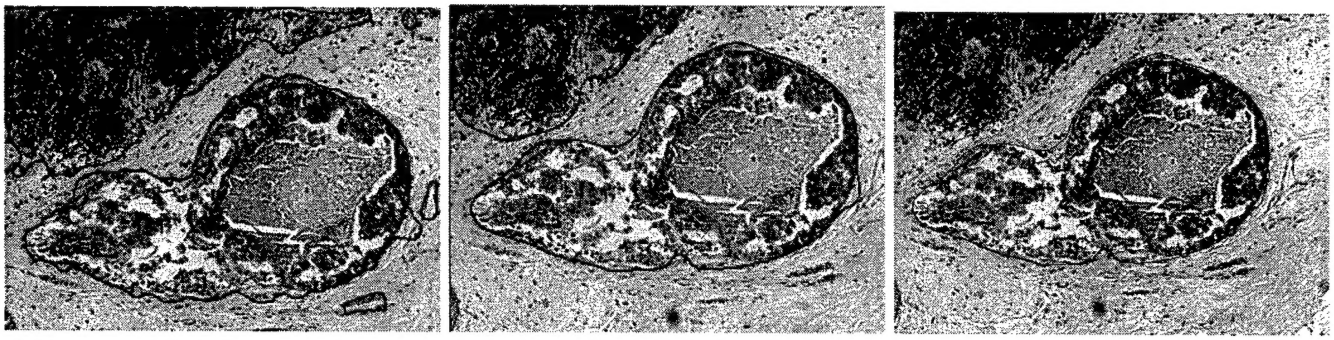


Fig. 6. Segmentation of a DCIS tumor in human mammary gland tissue. Left image: Initial contour (blue dot). Middle image: Results of the segmentation with red lines delineating tumor masses. Right image: results after edition using the interactive tools provided by the system.

Fig. 7. Segmentation of normal ducts in a mouse mammary gland tissue sample. Left image: initial data; middle image: results of the segmentation (red/black lines delineate normal ducts); right image: results after editing using the interactive tools provided by the system.

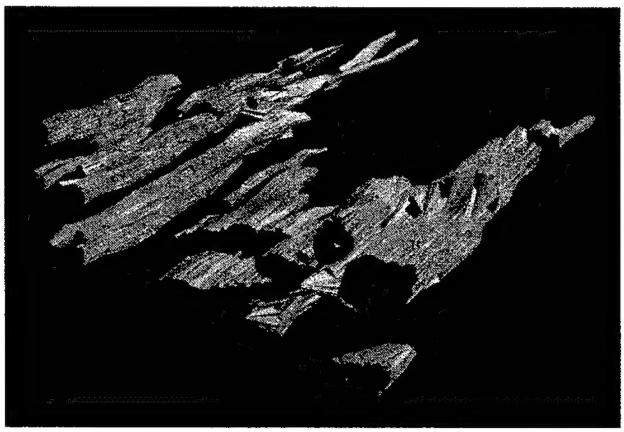


Fig. 8. 3D reconstruction of tumors in a human mammary gland tissue block. Tumor masses are rendered as gray volumes. The scene was stretched 10 times in the $Z$ direction to obtain a better view.

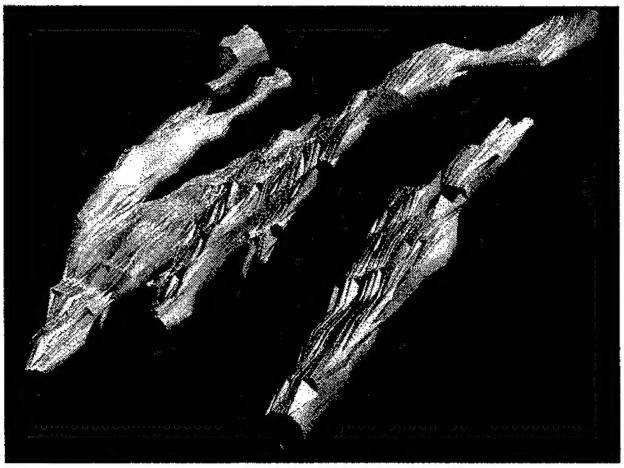


Fig. 9. 3D reconstruction of normal ducts in a mouse mammary gland. Ducts are rendered as gray volumes. A single duct and its branches can be traced throughout the gland. The Z direction was not stretched in this case.

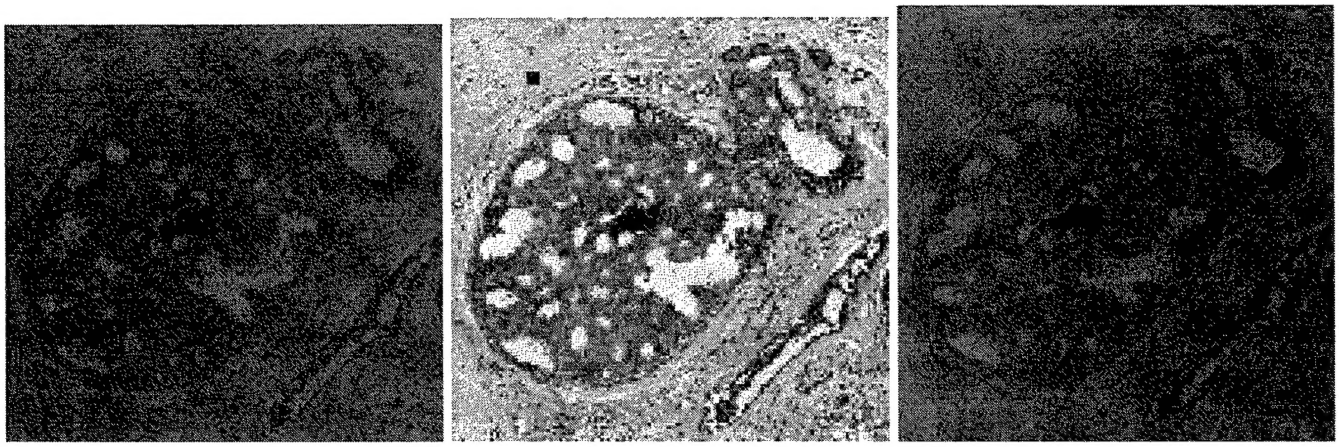


Fig. 10. Segmentation of a different DCIS tumor in human mammary gland tissue. Left image: the duct on the left contains a DCIS lesion with a necrotic center; the one on the right is normal. Middle image: the *ROI* was subsampled by a factor of 2 in both the X and Y directions; the blue dot represents the initial seed. Right image: results of the segmentation on the full resolution image.

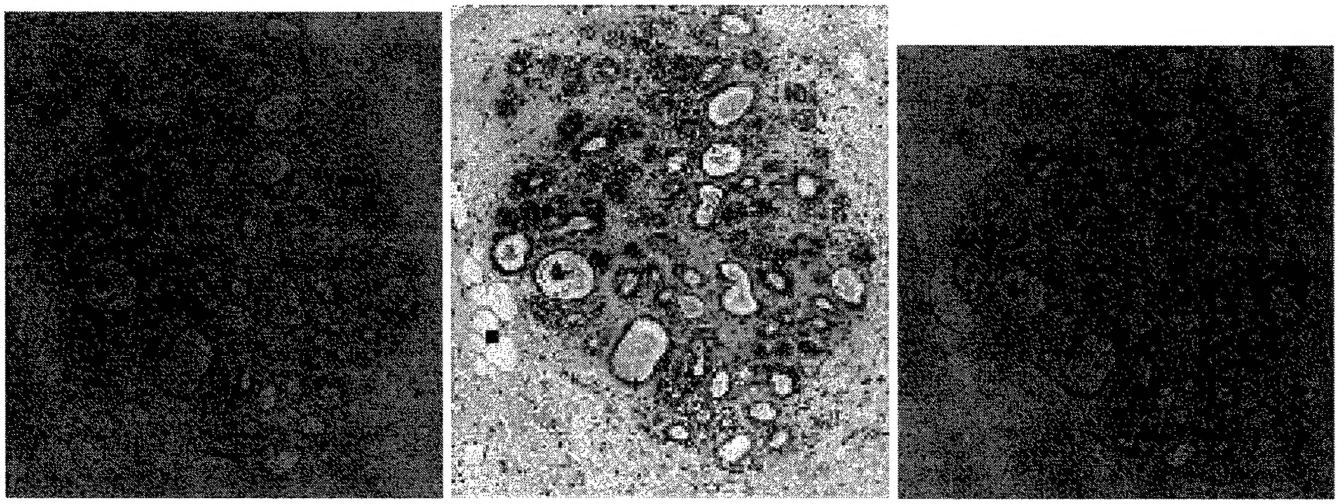


Fig. 11. Segmentation of a terminal ductal lobular unit in a human mammary gland. Left image: the duct on the left contains a section of a TDLU, one of the sites of milk production in the human breast. Middle image: the *ROI* was subsampled by a factor of 2 in both the X and Y directions; the blue dot represents the initial seed. Right image: results of the segmentation on the full resolution image.